PSEUDO TRANSIENT CONTINUATION AND TIME MARCHING METHODS FOR MONGE-AMPÈRE TYPE EQUATIONS

GERARD AWANOU

Abstract. We present two numerical methods for the fully nonlinear elliptic Monge-Ampère equation. The first is a pseudo transient continuation method and the second is a pure pseudo time marching method. The methods are proved to converge for smooth solutions. We give numerical evidence that they are also able to capture the viscosity solution of the Monge-Ampère equation. Even in the case of the degenerate Monge-Ampère equation, the time marching method appears also to compute the viscosity solution.

1. Introduction

We are interested in numerical solutions of the fully nonlinear Monge-Ampère equation

$$\det D^2 u = f \text{ in } \Omega, \quad u = g \text{ on } \partial\Omega, \tag{1.1}$$

on a convex bounded domain Ω of $\mathbb{R}^n, n = 2, 3$ with boundary $\partial\Omega$. The unknown u is a real valued function and f, g are given functions with $f > 0$ in the non degenerate case and $f \geq 0$ in the degenerate case. We will also assume that $f \in C(\Omega)$ and $g \in C(\partial\Omega)$. For $f \geq 0$ a solution is either convex or concave. It is therefore not restrictive to consider only convex solutions of (1.1).

Starting with [8, 16], interest has grown for finite element methods which are able to capture viscosity solutions of second order fully nonlinear equations. In the context of non-smooth solutions, for proven convergence results, wide stencils finite difference have been used for the Monge-Ampère and the Pucci equations [36]. See also [37] for a geometric approach. Outside of that framework, there are strong numerical evidence that standard finite difference methods [9] or standard finite element methods [33] converge as well to viscosity solutions. See also [20] and the references therein for the vanishing moment methodology. In this paper we give numerical evidence that C^1 conforming approximations of a variational formulation of (1.1) converge to viscosity solutions and new numerical evidence that standard finite difference methods converge as well. This is achieved by discretizing new iterative methods we introduce.

1.1. Contributions of this paper.

We establish in the context of C^1 conforming approximations that discrete functions near a strictly convex solution are strictly convex. This explains why convexity did not need to be imposed explicitly in some previous studies. Under the assumption that a discrete strictly convex solution exists, we prove the convergence of our iterative methods. In this paper, we prove the existence of a discrete strictly convex solution under the assumption that (1.1) has a smooth solution. In that case, Newton's method is the appropriate method for solving

the equation. We give a proof of convergence of Newton's method when the iterates are discretized with C^1 conforming approximations. Our numerical experiments indicate that in the degenerate case $f \geq 0$, our iterative methods enforce convexity for the two dimensional problem. We prove this result at the continuous level. In subsequent papers, we will prove existence and uniqueness of a discrete solution in the degenerate case, convergence to viscosity solutions, the convexity preserving property of our iterative methods at the discrete level as well as introduce new iterative methods, with proof of convergence, which enforce convexity in three dimensions.

The purpose of this paper is to introduce new iterative methods which are shown to numerically converge to viscosity solutions of (1.1) and to analyze their convergence for smooth solutions. Numerical results are given with the spline element method for which the author has extensive expertise. The finite difference results complements the ones obtained with the spline element method and illustrates the broad range of applicability of the methods presented.

To the author's best knowledge, this is the first time the pseudo transient method with $L = \Delta$ and the truncated time marching method have been shown to numerically converge to viscosity solutions of the Monge-Ampère equation.

As an initial guess for our iterative methods we take the solution of the Poisson equation $\Delta u = n f^{1/n}$ in $\Omega, u = g$ on $\partial\Omega$.

1.2. **Pseudo transient continuation method.** The first method is a pseudo transient continuation method. Given a fully nonlinear elliptic equation $F(u) = 0$ with F differentiable, we consider the sequence of problems

$$(\nu L + F'(u_k))(u_{k+1} - u_k) = -F(u_k),$$

where L is a linear operator which can be taken as the negative of the identity or $L = \Delta$ where Δ is the Laplace operator and $\nu > 0$ is a parameter. The motivation to consider pseudo transient continuation methods stems from the observation that Newton's method fails to converge to viscosity solutions. We refer to [29] for a review of these type of methods in the context of nonlinear equations. We presented the numerical results at the Thirteen International Conference on Approximation Theory in May 2010. In this paper, we will mainly consider the case where $L = \Delta$ which may be viewed as a preconditioner and in this case, the computations are faster compared to the case where L is the negative of the identity operator. Nethertheless, we will prove convergence of the methods at the discrete level for both cases. In the case of the Monge-Ampère equation, $F(u) = \det D^2 u - f$, the solution of the problem is reduced to solving a sequence of elliptic equations. Recall that in that case (see Lemma 2.3)

$$F'(u_k)(u_{k+1} - u_k) = \operatorname{div}\big((\operatorname{cof} D^2 u_k)(Du_{k+1} - Du_k) = (\operatorname{cof} D^2 u_k) : (D^2 u_{k+1} - D^2 u_k),$$

where $\operatorname{cof} D^2 u$ denotes the matrix of cofactors of the Hessian and $A : B$ denotes the Frobenius product of the matrices A and B.

Given an initial guess u_0, the pseudo transient method for the Monge-Ampère equation consists in solving the sequence of approximate problems

$$\nu L \theta_k + (\operatorname{cof}\, D^2 u_k) : D^2 \theta_k = (f - f_k), \quad f_k = \det D^2 u_k, \quad \theta_k = u_{k+1} - u_k. \tag{1.2}$$

1.3. **Pseudo time marching method.** The second method is a pseudo time marching method. Given $\nu > 0$, we consider the sequence of iterates

$$-\nu\Delta u_{k+1} = -\nu\Delta u_k + F(u_k),\ u_{k+1} = g \text{ on } \partial\Omega. \tag{1.3}$$

This can be interpreted as an Euler discretization of the pseudo time dependent equation $\frac{\partial \Delta u}{\partial t} + F(u) = 0$, or as a Laplacian preconditioner of a simple pseudo time marching algorithm, [26] $u_{k+1} = u_k - \frac{1}{\nu}\Delta^{-1}F(u_k)$. See also a remark in [34]. The simple pseudo time marching algorithm also performs well for numerical solutions in some cases for ν sufficiently large. However the use of the Laplacian preconditioner besides the speed of computation, also helps select a convex solution for the two dimensional Monge-Ampère equation.

The method we recommend for the degenerate Monge-Ampère equation (especially in three dimensions) is a truncated version of (1.3). For $m = 1, 2, \ldots$, we consider truncating functions $\chi_m(x)$ defined by $\chi_m(x) = -m$ for $x < -m$, $\chi_m(x) = x$ for $-m \leq x \leq m$ and $\chi_m(x) = m$ for $x > m$ and the sequence of problems

$$-\nu\Delta u^m_{k+1} = \chi_m(-\nu\Delta u^m_k + F(u^m_k)),\ u^m_{k+1} = g \text{ on } \partial\Omega. \tag{1.4}$$

We give numerical evidence that the limit of the discrete approximation of u^m_k as $k \to \infty$ and $m \to \infty$ approximates the viscosity solution of (1.1). Note that if F is smooth and the solution of (1.1) is also smooth, the right hand side of (1.4) is bounded and the method essentially reduces to (1.3). We have found (1.4) effective in the degenerate case $f \geq 0$ but $f > 0$ in Ω and in three dimensions with the finite difference method.

1.4. **Positivity preservation of the Laplacian.** We will see that the time marching method (with preconditioner) and the pseudo transient continuation method, enforce $\Delta u \geq 0$, which when combined with $\det D^2u = f \geq 0$ gives convexity for the two dimensional problem. We establish this result at the continuous level. The result at the continuous level explains but does not prove in the degenerate case why we also observe numerical convexity with the spline element method. Indeed a piecewise convex function which is globally C^1 is convex by [30] Lemma 1. Our existence and uniqueness result at the discrete level cover the cases $f \geq c_0 > 0$ for a constant c_0. In that case we show in the paper that convexity is automatically preserved in some neighborhood of the solution (see Lemma 2.5). The use of the spline element method is also motivated by its higher order of accuracy and its robustness in some limiting situations [3].

1.5. **Advantages and comparison of the two methods.** The advantage of (1.4) over the method of Oberman [36] is that standard finite difference methods can be used. Potentially, the use of wide stencils can be avoided. It was pointed out in [22] that the simple pseudo time marching algorithm suffers from a severe time-step restriction, is not suitable for solutions which are not strictly convex and does not enforce convexity. The truncated time marching method (1.4) numerically preserves convexity in two dimensions and for singular solutions appears to be more accurate than the pseudo transient methods for small values of the mesh size. Although the theory of the Monge-Ampère equation has concentrated on convex solutions, one can equally focus on concave solutions. We found out that (1.4) is better able to capture

concave solutions. It is easy to implement, requiring only a Poisson solver. For example one can capture weak solutions of the Monge-Ampere equation by simply discretizing (1.4) with the standard Lagrange finite elements. The time marching method can also be applied to fully nonlinear equations such as the Pucci equation where F is not differentiable. In summary the pseudo transient methods are better for smooth solutions and singular solutions on a coarse mesh. Otherwise the method of choice is the truncated time marching method.

The methods we propose can be used in the context of different types of discretizations allowing us in particular to treat more easily non-rectangular domains. The methods can be accelerated with fast Poisson solvers and multigrid methods. This latter property is even more striking for the time marching method as its implementation requires only having access to a multigrid Poisson solver.

1.6. **Newton's method.** It is believed that some methods are able to capture the viscosity solutions and others cannot. This paper gives evidence that at least in two dimensions, by relaxation and truncation, a method which works only for solutions in $H^2(\Omega)$ can be expected to perform in the non-smooth case for a class of problems which include the Monge-Ampère equation.

If one is only interested in smooth solutions of the Monge-Ampère equation, a number of methods have been recently proposed. In particular, Böhmer's method [10] has been recently implemented in [14] and Brenner et al [11] have used the efficient discontinuous Galerkin formulations. However, in all of the above cited work, Newton's method is used to solve the nonlinear system of equations resulting from the discretization. We postulate that solving the nonlinear equations with time marching methods captures weak solutions of the Monge-Ampère equation. In the case of conforming discretizations, this is evidenced by the numerical results of this paper. Since we are able to capture weak solutions with that technique with standard finite elements, the same should be true with the methods in [11]. Note that Newton's method has been used effectively in [24]. There, Newton's method is applied after the discretization of the equation. In our approach, the discretization process is applied on the Newton-Kantorovich iterations. Moreover in [24] a discretization which is known to converge to viscosity solutions is used on parts of the domain where the solution is not smooth. The term $\nu \Delta u$ in the methods we propose can be seen as an iterative regularization term. Pryer and Lakkis [33] have shown how to use Lagrange elements in a nonvariational formulation for the Monge-Ampère equation but their method has limited applications to more general fully nonlinear equations as they seem able to handle only the homogeneous Pucci equation.

1.7. **Main idea.** Our approach builds on the observation that globally C^1 spline approximations of a smooth strictly convex function remain strictly convex if the mesh is sufficiently fine, see Lemma 2.5 but also [30]. This allows us to transfer well established results for nonlinear elliptic equations, e.g. [19], to the context of smooth solutions of the elliptic Monge-Ampère equation. In this paper, we prove convergence results for the iterative methods introduced above from that point of view.

1.8. **Organization of the paper.** The paper is organized as follows: in the second section, we first prove the convergence of the methods in Hölder spaces and show how they preserve positivity of the Laplacian and hence convexity in two dimensions. We then study the convergence in Sobolev spaces for smooth solutions in the third section. After some preliminaries, we show that a discrete strictly convex solution exists when (1.1) has a solution which is C^2 up to the boundary. We give error estimates and convergence of Newton's method. The proof of convergence of our iterative methods conclude the third section. The last section is devoted to numerical results with the spline element method and finite difference methods. It includes also a brief description of the spline element method.

2. Convergence of the iterative methods

In this section, we first describe the convergence properties of the methods in Hölder spaces and establish that they preserve the positivity of the Laplacian. We then give convergence results at the discrete level for smooth solutions in Sobolev spaces.

2.1. **Convergence of pseudo transient continuation methods in Hölder spaces.** We consider a damped version of (1.2), namely

$$\begin{aligned} &(\nu L + F'(u_k))(u_{k+1} - u_k) = -\frac{1}{\tau} F(u_k), \\ &\nu L\theta_k + (\text{cof } D^2u_k) : D^2\theta_k = \frac{1}{\tau}(f - f_k), \quad f_k = \det D^2u_k, \quad \theta_k = u_{k+1} - u_k, \end{aligned} \tag{2.1}$$

where $\tau > 0$ is a damping parameter. In the numerical experiments, we used $\tau = 1$. We will need the following global regularity result, [40].

Theorem 2.1. *Let Ω be a uniformly convex domain in $\mathbb{R}^n$, with boundary in C^3. Suppose $g \in C^3(\overline{\Omega})$, inf $f > 0$, and $f \in C^\alpha$ for some $\alpha \in (0, 1)$. Then (1.1) has a convex solution u which satisfies the a priori estimate*

$$||u||_{C^{2,\alpha}(\overline{\Omega})} \leq C,$$

where C depends only on n, α, inf $f, \Omega, ||f||_{C^\alpha(\overline{\Omega})}$ and $||g||_{C^3}$.

According to [40], all assumptions in the above theorem are sharp. We have the following analogue of Theorem 2.1 in [34].

Theorem 2.2. *Let Ω be a uniformly convex domain in $\mathbb{R}^n$, with boundary in C^3. Let $0 < m \leq f \leq M, f \in C^\alpha$ for some $m, M > 0$ and $\alpha \in (0, 1)$. Assume also that $g \in C^3(\overline{\Omega})$. Then, there exists $\tau \geq 1$ depending on m, f, such that if u_k is the sequence defined by (2.1), it converges in $C^{2,\beta}$ to a solution u of (1.1) for u_0 sufficiently close to u and for every $\beta < \alpha$.*

Proof. The proof follows essentially [34]. One shows by induction that there are constants $C_1, C_2 > 0$ such that

$$\frac{1}{C_1} \det D^2u \leq \det D^2u_k \leq C_1 \det D^2u \text{ and } ||\det D^2u - \det D^2u_k||_{C^\alpha} \leq C_2, \tag{2.2}$$

for τ sufficiently large and ν in an appropriate range. This implies that the sequence $f_k = \det D^2u_k$ is bounded in C^α and by Theorem 2.1, the sequence u_k is bounded in

$C^{2,\alpha}$. Arzela-Ascoli's theorem is then used to prove that the sequence is precompact in $C^{2,\beta}, \beta < \alpha$. Since (1.1) has at most two solutions, [13] p. 324, by requiring u_0 sufficiently close to u, we assure that the solution is locally unique. □

We now explain how the condition $\Delta u \geq 0$ is preserved in the iteration

$$\nu \Delta u_{k+1} + (\text{cof } D^2 u_k) : D^2 u_{k+1} = \nu \Delta u_k + (\text{cof } D^2 u_k) : D^2 u_k - \det D^2 u_k + f. \tag{2.3}$$

Recall that the eigenvalues are continuous functions of the entries of $D^2 v$ as roots of the characteristic equation, [38] Appendix K, or [27]. Hence for $0 < m \leq f \leq M$ as in the theorem, v is strictly convex whenever $||u-v||_{C^{2,\beta}(\Omega)}$ is sufficiently small. Since the sequence u_k was shown to converge in $C^{2,\beta}(\Omega), \beta < \alpha$, convexity is automatically enforced not only for u but for u_k for k sufficiently large.

Now assume that $f \geq 0$ and that the sequence u_k has been shown to converge to u in $C^{2,\beta}(\Omega)$ for some β in $(0,1)$. From the arithmetic-geometric inequality, we have

$$\frac{(\Delta u_k)^n}{n^n} \geq \det D^2 u_k.$$

Again, by the continuity of the eigenvalues, Δv is bounded in a neighborhood of u in which all u_k belong for k large enough. Choose ν such that $\nu \geq (n-1)(\Delta u_k)^{n-1}/n^n$ for all k and note that the right hand of (2.3) is equal to $\nu \Delta u_k + (n-1) \det D^2 u_k + f$. By the assumption on ν, we get $\nu \Delta u_{k+1} + (\text{cof } D^2 u_k) : D^2 u_{k+1} \geq 0$. In the limit, we obtain $\nu \Delta u + (\text{cof } D^2 u) : D^2 u \geq 0$. Since $\det D^2 u \geq 0$ by assumption, we get $\Delta u \geq 0$.

As for the time marching method

$$-\nu \Delta u_{k+1} = -\nu \Delta u_k + \det D^2 u_k - f, \ u_{k+1} = g \text{ on } \partial \Omega,$$

assume now again that $f \geq 0$ and that the sequence u_k has been shown to converge to u in $C^{2,\beta}(\Omega)$ for some β in $(0,1)$. Choose ν such that $\nu \geq (\Delta u_k)^{n-1}/n^n$. We have $-\nu \Delta u_{k+1} = -\nu \Delta u_k + \det D^2 u_k - f$. It follows from the arithmetic-geometric inequality that

$$\nu \Delta u_k \geq \frac{(\Delta u_k)^n}{n^n} \geq \det D^2 u_k.$$

and so $-\nu \Delta u_k + \det D^2 u_k \leq 0$ and it follows that the time marching method also preserves the positivity of the Laplacian.

2.2. **Convergence at the discrete level in Sobolev spaces.** We first give a variational formulation of (1.1) and its discretization by conforming finite dimensional spaces. Then we prove a number of lemmas which are needed for our convergence proofs. In particular we establish in Lemma 2.5 that discrete convex functions near a strictly convex solution are strictly convex. We also show the existence of a strictly discrete convex solution when the equation has a strictly convex smooth solution and give error estimates. We prove convergence of Newton's method and of our iterative methods.

We use the standard notation of Sobolev spaces $W^{k,p}(\Omega)$ with norms $||.||_{k,p}$ and semi-norm $|.|_{k,p}$. In particular, $H^k(\Omega) = W^{k,2}(\Omega)$ and in this case, the norm and semi-norms will be denoted respectively by $||.||_k$ and semi-norm $|.|_k$. We make the usual convention of denoting constants by C but will occasionally index some constants. For constants

which depend on the mesh size, we use $c(h)$ or $C(h)$. We make the assumption that the boundary of Ω is polygonal and that the triangulation is shape regular in the sense that there is a constant $C > 0$ such that any triangle K, $h_K/\rho_K \leq C$, where h_K denotes the diameter of K and ρ_K the radius of the largest ball contained in K. We also require the triangulation to be quasi-uniform in the sense that h/h_{min} is bounded where h and h_{min} are the maximum and minimum respectively of $\{h_K, K \in \mathcal{T}_h\}$.

We choose

$$V^h := S^r_d(\mathcal{T}) = \{s \in C^r(\Omega),\ s|_t \in \mathcal{P}_d,\ \forall t \in \mathcal{T}\}. \tag{2.4}$$

for a triangulation $\mathcal{T}$ of the domain and $\mathcal{P}_d$ the space of polynomials of degree less than or equal to d. In two dimensions, it is known that, [32], for $d \geq 3r+2$ and $0 \leq l \leq d$, there exists a linear quasi-interpolation operator Q mapping $L_1(\Omega)$ into the spline space $S^r_d(\mathcal{T})$ and a constant C such that if f is in the Sobolev space $W^{l+1,p}(\Omega), 1 \leq p \leq \infty$

$$||f - Qf||_{k,p} \leq Ch^{l+1-k}|f|_{l+1,p}, \tag{2.5}$$

for $0 \leq k \leq l$. If Ω is convex, the constant C depends only on d, l and on the smallest angle θ_h in $\mathcal{T}$. In the nonconvex case, C depends only on the Lipschitz constant associated with the boundary of Ω.

In three dimensions, the result holds for $d \geq 8r+1$, c.f. [32]. It is also known c.f. [17] that the full approximation property for spline spaces holds for certain combinations of d and r on special triangulations. Note that, by (2.5),

$$||Qu||_{2,p} \leq C||u||_{2,p}, \quad u \in W^{2,p}(\Omega), \tag{2.6}$$

for all p. We assume that $r \geq 1$ and that the following inverse inequality which holds for finite element spaces, c.f. Theorem 4.5.11 of [12], also holds for the spline spaces

$$||u||_{2,p} \leq Ch^{l-2+\min(0,\frac{n}{p}-\frac{n}{q})}||u||_{l,q}, \forall u \in V^h, \tag{2.7}$$

for $0 \leq l \leq 2, 1 \leq p, q \leq \infty$. The local estimates may be viewed as a consequence of the assumption of uniform triangulation and of Markov inequality, [32] p. 2. Passing from local estimates to global estimates can be done as in [31].

The variational formulation of (1.1) is given by: find $u \in W^{2,n}(\Omega)$, $u = g$ on $\partial\Omega$ such that

$$-\frac{1}{n}\int_\Omega (\operatorname{cof} D^2u)Du \cdot Dw \ dx = \int_\Omega fw \ dx, \quad \forall w \in W^{2,n}(\Omega) \cap H^1_0(\Omega). \tag{2.8}$$

To see that the left hand side is bounded for $u \in W^{2,n}(\Omega)$, notice that for $n = 2$

$$\left| \int_\Omega (\operatorname{cof} D^2u)Du \cdot Dw \ dx \right| \leq C||D^2u||_{0,2}||Du||_{0,4}||Dw||_{0,4}.$$

Next for $u \in H^2(\Omega), \partial u/\partial x_i \in H^1(\Omega), i = 1, \ldots, n$ and by the embedding of $H^1(\Omega)$ in $L^q(\Omega)$ for $q \geq 1$ when $n = 2$, the right hand side above is bounded by $C||D^2u||_{L^2(\Omega)}||u||_{H^2(\Omega)}$ $||w||_{H^2(\Omega)}$. In three dimensions, the term $\operatorname{cof} D^2u$ involves the product of two second order derivatives. We have by Hölder's inequality and the embedding of $H^1(\Omega)$ in

$L^q(\Omega)$ for $1 \leq q \leq 6$ when $n=3$,

$$\begin{aligned}\Big| \int_\Omega \frac{\partial^2 u}{\partial x^2}\frac{\partial^2 u}{\partial z^2}\frac{\partial u}{\partial x}\frac{\partial u}{\partial y}\, dx\Big| &\leq ||\frac{\partial^2 w}{\partial x^2}||_{0,3}||\frac{\partial^2 u}{\partial x^2}||_{0,3}||\frac{\partial^2 u}{\partial x^2}||_{0,6}||\frac{\partial^2 w}{\partial x^2}||_{0,6} \\ &\leq ||u||_{2,3}^2||u||_2||w||_2.\end{aligned}$$

We conclude that for $n=3$,

$$\Big| \int_\Omega (\operatorname{cof} D^2 u) Du \cdot Dw \ dx \Big| \leq C||u||_{2,3}^2||u||_2||w||_2.$$

We note that for $n=2,3$, we may write

$$\Big| \int_\Omega (\operatorname{cof} D^2 u) Du \cdot Dw \ dx \Big| \leq C||u||_{2,n}^{n-1}||u||_2||w||_2. \tag{2.9}$$

Put $V = W^{2,n}(\Omega)$ and $V_0 = W^{2,n}(\Omega) \cap H_0^1(\Omega)$. Let V^h be a conforming finite dimensional subspace of $W^{2,n}(\Omega)$, V_0^h be a conforming finite dimensional subspace of $W^{2,n}(\Omega) \cap H_0^1(\Omega)$. Furthermore let g_h be the interpolant in V^h of a continuous extension of g. We have the following conforming discretization of (2.8): find $u_h \in V^h$, $u_h = g_h$ on $\partial\Omega$ such that

$$-\frac{1}{n}\int_\Omega (\operatorname{cof} D^2 u_h) Du_h \cdot Dw_h \ dx = \int_\Omega f w_h \ dx, \quad \forall w_h \in V_0^h. \tag{2.10}$$

We now prove a number of preliminary results. The following lemma is elementary.

Lemma 2.3. *We have*

$$\det D^2 u = \frac{1}{n}(cof D^2 u) : D^2 u = \frac{1}{n} \operatorname{div} \big((cof D^2 u) Du \big). \tag{2.11}$$

And for $F(u) = \det D^2 u$ we have

$$F'(u)(w) = (cof D^2 u) : D^2 w = \operatorname{div} \big((cof D^2 u) Dw \big),$$

for u, w sufficiently smooth.

Proof. Note that for any $n \times n$ matrix A, $\det A = 1/n(\operatorname{cof} A) : A$, where $\operatorname{cof} A$ is the matrix of cofactors of A and for two $n \times n$ matrices M, N, $M : N = \sum_{i,j=1}^n M_{ij} N_{ij}$. For any sufficiently smooth matrix field A and vector field v, $\operatorname{div} A^T v = (\operatorname{div} A) \cdot v + A : Dv$. Here the divergence of a matrix field is the divergence operator applied row-wise. If we put $v = Du$, then $\det D^2 u = 1/n(\operatorname{cof} D^2 u) : D^2 u = 1/n\, (\operatorname{cof} Dv) : Dv$ and $\operatorname{div}(\operatorname{cof} Dv)^T v = \operatorname{div}(\operatorname{cof} Dv) \cdot v + (\operatorname{cof} Dv) : Dv$. But $\operatorname{div} \operatorname{cof} Dv = 0$, c.f. for example [18] p. 440. Hence since $D^2 u$ and $\operatorname{cof} D^2 u$ are symmetric matrices (2.11) follows. The assertion about the Fréchet derivative of F also follows from these identities. □

Recall that $W^{2,p}(\Omega)$ is continuously embedded in $H^2(\Omega)$ for $p \geq 2$ and hence for $u \in W^{2,n}(\Omega), n \geq 2$

$$||u||_2 \leq C||u||_{2,n}. \tag{2.12}$$

Lemma 2.4. *Let $v, w \in W^{2,n}(\Omega), n = 2, 3$ and $\psi \in H_0^1(\Omega) \cap H^2(\Omega)$, then*

$$\int_\Omega (\det D^2 v - \det D^2 w)\psi \, dx = -\int_0^1 \left\{ \int_\Omega \big((\text{cof}[(1-t)D^2 w + tD^2 v)] \right. \\ \left. (Dv - Dw)\big) \cdot D\psi \, dx \right\} dt,$$

and

$$\left| \int_\Omega (\det D^2 v - \det D^2 w)\psi \, dx \right| \leq C_0(||v||_{2,n} + ||w||_{2,n})^{n-1} ||v - w||_2 ||\psi||_2. \tag{2.13}$$

Moreover, if $v, w \in W^{2,n}(\Omega) \cap W^{2,\infty}(\Omega)$

$$\left| \int_\Omega (\det D^2 v - \det D^2 w)\psi \, dx \right| \leq C_0(||v||_{2,\infty} + ||w||_{2,\infty})^{n-1} ||v - w||_1 ||\psi||_1. \tag{2.14}$$

Proof. We first recall the Mean Value Theorem. Let E and F be Banach spaces and and let us denote by $L(E, F)$ the space of continuous linear mappings from E to F. Let also X be an open subset of E and let $F : X \to F$ be a differentiable map. If $F' : X \to L(E, F)$ is continuous, F is said to be of class C^1 and for all $a, x \in X$, we have

$$F(x) = F(a) + \int_0^1 F'[(1-t)a + tx](x - a) \, dt.$$

Next, let $F : C^\infty(\Omega) \to C^\infty(\Omega)$ denote the mapping $v \mapsto \det D^2 v$. Then F is differentiable with

$$F'[u](v) = (\text{cof}\, D^2 u) : D^2 v = \text{div}\, \big((\text{cof}\, D^2 u) Dv\big).$$

Since $v \mapsto F'[v]$ is linear, F is of class C^1 and by the Mean Value Theorem

$$F(v) - F(w) = \int_0^1 \text{div}\, \big((\text{cof}\,(1-t)D^2 w + tD^2 v)(Dv - Dw)\big) \, dt.$$

Next, let $\psi \in \mathcal{D}(\Omega)$. By Fubini's theorem,

$$\begin{aligned} \int_\Omega (\det D^2 v - \det D^2 w)\psi \, dx &= \int_0^1 \left\{ \int_\Omega \text{div}\, \big((\text{cof}\,(1-t)D^2 w + tD^2 v)(Dv - Dw)\big)\psi \, dx \right\} dt \\ &= -\int_0^1 \left\{ \int_\Omega \big((\text{cof}\,(1-t)D^2 w + tD^2 v)(Dv - Dw)\big) \cdot D\psi \, dx \right\} dt. \end{aligned}$$

As in the proof of (2.9), we conclude that (2.13) holds for u, ψ smooth, $\psi = 0$ on $\partial\Omega$. Then by the density of $\mathcal{D}(\Omega)$ in $H_0^1(\Omega)$ it also holds for $v, w \in C^\infty(\Omega)$ and $\psi \in H_0^1(\Omega) \cap H^2(\Omega)$. Finally by the density of $\mathcal{D}(\Omega)$ in the Sobolev spaces, it is also valid for $v, w \in W^{2,n}(\Omega)$. Inequality (2.14) is proved similarly. □

The following lemma is fundamental.

Lemma 2.5. *Let u be a $C^2(\Omega)$ strictly convex function. Then there exists $\delta > 0$ such that for $||v - u||_{2,\infty} \leq \delta$, v is strictly convex. Moreover if $u \in W^{2,\infty}(\Omega) \cap H^{l+1}(\Omega), 1+$*

$n/2 < l \leq d$, there exists $\delta(h) > 0$ such that for $v_h \in V^h, ||v_h - u||_1 \leq \delta(h)$ and h sufficiently small, v_h is strictly convex and

$$m||w||_1^2 \leq \int_\Omega [(\operatorname{cof} D^2 v_h(x)) Dw(x)] \cdot Dw(x)\ dx \leq M||w||_1^2, w \in H_0^1(\Omega), \tag{2.15}$$

for constants $m, M > 0$ which are independent of h. It follows that for h sufficiently small, $r \geq 1$, Qu is strictly convex.

Proof. Let $\lambda_1(D^2u(x))$ and $\lambda_2(D^2u(x))$ be the smallest and largest eigenvalues of $D^2u(x)$ respectively. Since u is strictly convex, we have

$$2m \leq \lambda_1(D^2u(x)) \leq \lambda_2(D^2u(x)) \leq M/2, \text{ for } m, M > 0.$$

Since the eigenvalues are continuous functions of the entries of the Hessian, as roots of the characteristic equation, for all $\epsilon > 0$, there exists $\delta > 0$ such that $||v-u||_{C^2(\Omega)} \leq \delta$ implies $|\lambda_i(D^2v(x)) - \lambda_i(D^2u(x))| < \epsilon, i = 1, 2, \forall x \in \Omega$. Since $\lambda_1(D^2u(x)) > 2m, \forall x \in \Omega$, for $\epsilon = m$, we get $\lambda_1(D^2v(x)) > m, \forall x \in \Omega$. We conclude that for $||v-u||_{C^2(\Omega)} \leq \delta$, $\lambda_1(D^2v(x)) > m, \forall x \in \Omega$. Similarly, $\lambda_2(D^2v(x)) < M, \forall x \in \Omega$. This implies that v is strictly convex for $||v - u||_{C^2(\Omega)} \leq \delta$.

Since $\lambda_1(D^2v(x))$ and $\lambda_2(D^2v(x))$ are the minimum and maximum of the Rayleigh quotient $(\operatorname{cof} D^2v(x)z) \cdot z/||z||^2$, where $||z||$ denotes the standard Euclidean norm in $\mathbb{R}^n$, we have

$$m||z||^2 \leq [(\operatorname{cof} D^2v(x))z] \cdot z \leq M||z||^2, z \in \mathbb{R}^n.$$

This implies

$$m||w||_1^2 \leq \int_\Omega [(\operatorname{cof} D^2 v(x)) Dw(x)] \cdot Dw(x)\ dx \leq M||w||_1^2, w \in H_0^1(\Omega), \tag{2.16}$$

where we used the equivalence of the seminorm $|\,|_1$ and norm $||\,||_1$ on $H_0^1(\Omega)$ Next, on each element K, there is $\delta(K) > 0$ such that for $||v_h - u||_{C^2(K)} \leq \delta(K)$, (2.16) holds with the domain Ω replaced by K. Let then $\delta(h)$ be the minimum of $\{\delta(K), K$ in $\mathcal{T}_h\}$. Since by (2.5) and (2.7), for $l \geq 2$

$$\begin{aligned}
||v_h - u||_{2,\infty} &\leq ||v_h - Qu||_{2,\infty} + ||Qu - u||_{2,\infty} \\
&\leq \frac{C}{h^{1+\frac{n}{2}}}||v_h - Qu||_1 + Ch^{l-1}|u|_{l+1,\infty} \\
&\leq \frac{C}{h^{1+\frac{n}{2}}}||v_h - u||_1 + \frac{C}{h^{1+\frac{n}{2}}}||u - Qu||_1 + Ch^{l-1}|u|_{l+1,\infty} \\
&\leq \frac{C}{h^{1+\frac{n}{2}}}||v_h - u||_1 + Ch^{l-1-\frac{n}{2}}|u|_{l+1} + Ch^{l-1}|u|_{l+1,\infty},
\end{aligned}$$

we see that if $||v_h - u||_1 < h^{1+n/2}\delta(h)/(3C)$ and h sufficiently small, $||v_h - u||_{C^2(K)} \leq ||v_h - u||_{2,\infty} \leq \delta(h)$ and v_h is piecewise convex. Since $r \geq 1$, v_h is of class C^1 and piecewise convex that is, v_h is convex by [30], Lemma 1.

Since by (2.9), $\int_\Omega [(\operatorname{cof} D^2v_h(x))Dw(x)] \cdot Dw(x)\ dx$ is finite for $w \in H_0^1(\Omega)$ and $v_h \in V^h \subset W^{2,n}(\Omega)$, we conclude that (2.15) holds.

□

2.2.1. *Existence, uniqueness and error estimates for C^1 conforming approximations.* We show in this section that if (1.1) has a smooth strictly convex solution, then the discrete equations (2.10) also have a unique strictly convex discrete solution. The existence and uniqueness of the solution of (2.10) was also addressed by Böhmer [10] from a different perspective. Böhmer's method requires stable splitting of spline spaces [15] and the required modification of the Argyris space has been implemented only recently [14]. We adapt and simplify the methods of [21] which were given for the vanishing moment formulation. In addition to existence and uniqueness, this also gives error estimates. See also [35] for another discussion of conforming approximation for homogeneous boundary conditions. The latter study has been completed after the results of this work were announced. We note that one may be able to prove the existence of a solution at the discrete level following [39] as indicated in the proof of Theorem 5.8 of [19].

Lemma 2.6. *Assume $u \in W^{l+1,\infty}(\Omega), 3 \leq l \leq d$ is a strictly convex function, that Ω is convex with a Lipschitz continuous boundary and that the spaces $V^h = S^r_d(\mathcal{T})$ have the optimal approximation property* (2.5) *and satisfy the inverse estimates* (2.7)*. Then the problem* (2.10) *has a unique strictly convex solution u_h and we have the error estimates*

$$||u-u_h||_2 \leq Ch^{l-1}|u|_{l+1}, \ ||u-u_h||_1 \leq Ch^l|u|_{l+1}$$
$$||u-u_h||_{L^2} = O(h^{l+1}) \text{ for } h \text{ sufficiently small.}$$

Proof. We consider the linear problem: find $v \in H^1_0(\Omega)$ such that

$$\int_\Omega (\operatorname{cof} D^2u) Dv \cdot Dw \ dx = \int_\Omega \phi\, w \ dx, \forall w \in H^1_0(\Omega), \tag{2.17}$$

for $\phi \in L^2(\Omega)$. Since $u \in W^{l+1,\infty}(\Omega), 3 \leq l \leq d$ is strictly convex, $\exists\, m, M > 0$ such that

$$m||w||_1^2 \leq \int_\Omega [(\operatorname{cof} D^2u(x))Dw(x)] \cdot Dw(x) \ dx \leq M||w||_1^2, w \in H^1_0(\Omega).$$

The existence and uniqueness of a solution to (2.17) follows immediately from Lax-Milgram lemma. Similarly, there exists a unique solution to the problem: Find $v_h \in V^h_0$ such that

$$\int_\Omega (\operatorname{cof} D^2u) Dv_h \cdot Dw_h \ dx = \int_\Omega \phi\, w_h \ dx, \forall w_h \in V^h_0. \tag{2.18}$$

We note that the constant C above depends on u and that since Ω is assumed convex, $v \in H^2(\Omega)$ by elliptic regularity.

We define a bilinear form on $H^1_0(\Omega) \times H^1_0(\Omega)$ by

$$B[v,w] = \int_\Omega (\operatorname{cof} D^2u) Dv \cdot Dw \ dx, \tag{2.19}$$

and for a given $v_h \in V^h$, $v_h = g_h$ on $\partial\Omega$, define $T(v_h)$ as the unique solution of

$$B[v_h - T(v_h), \psi_h] = -\int_\Omega \det D^2 v_h \psi_h \ dx + \int_\Omega f \psi_h \ dx, \quad \forall \psi_h \in V^h_0. \tag{2.20}$$

Since $v_h - T(v_h) \in V_0^h$, $T(v_h) \in V^h$, $T(v_h) = g_h$ on $\partial\Omega$. A fixed point of the nonlinear operator T corresponds to a solution of (2.10) and conversely if v_h is a solution of (2.10), then v_h is a fixed point of T. We will show that T has a unique fixed point in a neighborhood of $Q(u)$. Put

$$B_h(\rho) = \{v_h \in V^h, v_h = g_h \,\text{on}\, \partial\Omega, ||v_h - Qu||_1 \leq \rho\}.$$

We first show that

$$||Qu - T(Qu)||_1 \leq C_1 h^l ||u||_{2,\infty}^{n-1} |u|_{l+1}, \tag{2.21}$$

then we show there exists $0 < \rho_0$ depending on h such that T is a contraction mapping in the ball $B_h(\rho_0)$ with a contraction factor $1/2$. We conclude by applying the Brouwer fixed point theorem in a suitable ball.

Put $w_h = Qu - T(Qu)$. Then $w_h \in H_0^1(\Omega)$ and using $\det D^2 u = f$,

$$B[w_h, w_h] = \int_\Omega \big(\det D^2 u - \det D^2 Qu\big) w_h \, dx.$$

Then by Lemma 2.4, the coercivity of B on $H_0^1(\Omega)$, (2.6) and (2.5),

$$||w_h||_1^2 \leq C ||u||_{2,\infty}^{n-1} ||u - Qu||_1 ||w_h||_1 \leq C_1 h^l ||u||_{2,\infty}^{n-1} |u|_{l+1} ||w_h||_1,$$

from which (2.21) follows.

For $v_h, w_h \in B_h(\rho_0)$, with ρ_0 yet to be determined, and $\psi_h \in V_0^h$,

$$\begin{aligned} B[T(v_h) - T(w_h), \psi_h] &= B[T(v_h) - v_h, \psi_h] + B[v_h - w_h, \psi_h] + B[w_h - T(w_h), \psi_h] \\ &= \int_\Omega (\det D^2 v_h - \det D^2 w_h) \psi_h \, dx \\ &\qquad + \int_\Omega [(\text{cof})\, D^2 u (Dv_h - Dw_h) \ \cdot D\psi_h \, dx. \end{aligned}$$

Using

$$\int_\Omega \text{cof}\, D^2 u (Dv_h - Dw_h) D\psi_h \, dx = \int_0^1 \int_\Omega [(\text{cof}\, D^2 u)(Dv_h - Dw_h)] \cdot D\psi_h \, dx \, dt,$$

Lemma 2.4, and with the observation that by the inverse inequality, $v_h, w_h \in H^2(\Omega)$, we have

$$\begin{aligned} B[T(v_h) - T(w_h), \psi_h] &= \int_0^1 \bigg\{ \int_\Omega [(\text{cof}\, D^2 u - \text{cof}\,((1-t) D^2 w_h + t D^2 v_h)) \\ &\qquad\qquad (Dv_h - Dw_h)] \cdot D\psi_h \, dx \bigg\} \, dt \\ &= \int_0^1 \bigg\{ \int_\Omega [(\text{cof}\, D^2 u - \text{cof}\, D^2 Qu)(Dv_h - Dw_h)] \cdot D\psi_h \, dx \bigg\} \, dt \\ &\quad - \int_0^1 \bigg\{ \int_\Omega (\text{cof}\,[((1-t)(D^2 w_h - D^2 Qu) + t(D^2 v_h - D^2 Qu)) \\ &\qquad\qquad (Dv_h - Dw_h)] \cdot D\psi_h \, dx \bigg\} \, dt. \end{aligned}$$

We conclude using $\psi_h = T(v_h) - T(w_h)$, the coercivity of B on $H_0^1(\Omega)$, and (2.9), that

$$||T(v_h) - T(w_h)||_1^2 \leq C||u - Qu||_{2,\infty}^{n-1}||v_h - w_h||_1||\psi_h||_1 \\ + C(||w_h - Qu||_{2,n}^{n-1} + ||v_h - Qu||_{2,n}^{n-1})||v_h - w_h||_2||\psi_h||_2.$$

For $n = 2$, by the inverse inequality (2.7) and (2.5),

$$||T(v_h) - T(w_h)||_1 \leq (C_2 h^{l-1}|u|_{l+1,\infty} + C_3\frac{\rho_0}{h^3})||v_h - w_h||_1.$$

We require $C_2 h^{l-1}|u|_{l+1,\infty} \leq 1/4$ and we choose ρ_0 such that, $C_3\rho_0/h^3 \leq 1/4$ for example $\rho_0 = h^3/(4C_3)$. It follows that T is a contraction mapping in the ball $B_h(\rho_0)$ with a contraction factor $1/2$. We now consider the case $n = 3$. We have by (2.7)

$$||T(v_h) - T(w_h)||_1 \leq Ch^{2(l-1)}|u|_{l+1,\infty}^2||v_h - w_h||_1 \\ + C(||w_h - Qu||_{2,3}^2 + ||v_h - Qu||_{2,3}^2)\frac{||v_h - w_h||_1}{h^2}.$$

Again, by (2.7),

$$||v||_{2,3} \leq \frac{C}{h^{\frac{1}{2}}}||v||_2 \leq \frac{C}{h^{\frac{3}{2}}}||v||_1.$$

We conclude that for $n = 3$,

$$||T(v_h) - T(w_h)||_1 \leq (C_4 h^{2(l-1)}|u|_{l+1,\infty}^2 + C_5\frac{\rho_0^2}{h^5})||v_h - w_h||_1.$$

Now, we require $C_4 h^{2(l-1)}|u|_{l+1,\infty}^2 < 1/4$ and choose ρ_0 such that $C_5\rho_0^2/h^5 < 1/4$. Finally, note that with ρ_0 sufficiently small, in $B_h(\rho_0)$,

$$||Tv_h - Qu||_1 \leq ||Qu - T(Qu)||_1 + ||TQu - Tv_h||_1 \leq C_1 h^l||u||_{2,\infty}^{n-1}|u|_{l+1} + \frac{||v_h - Qu||_1}{2}.$$

Put $\rho_1 = 2C_1 h^l||u||_{2,\infty}^{n-1}|u|_{l+1}$. Since $l \geq 3$ and h sufficiently small, $\rho_1 \leq \rho_0$, and T maps $B_h(\rho_1)$ into itself.

We conclude by the Brouwer fixed point theorem that T has a unique fixed point u_h in $B_h(\rho_1)$ to which the iterates $u_h^{k+1} = T(u_h^k)$ converge. Moreover

$$||u - u_h||_1 \leq ||u - Qu||_1 + ||Qu - u_h||_1 = ||u - Qu||_1 + ||Qu - T(u_h)||_1 \\ \leq Ch^l|u|_{l+1} + \rho_1 \leq (C + 2C_1||u||_{2,\infty}^{n-1})h^l|u|_{l+1}.$$

and

$$||u - u_h||_2 \leq ||u - Qu||_2 + ||Qu - u_h||_2 = ||u - Qu||_2 + ||Qu - Tu_h||_2 \\ \leq Ch^{l-1}|u|_{l+1} + \frac{\rho_1}{h} \leq Ch^{l-1}|u|_{l+1},$$

using again an inverse estimate.

Using a duality argument, we prove the L^2 error estimate. Recall that the domain is convex and let $w \in H^2(\Omega)$ be the solution of the problem

$$\operatorname{div}(\operatorname{cof} D^2u)Dw = u - u_h, \text{ in } \Omega, w = 0 \text{ on } \Omega.$$

By elliptic regularity,

$$||w||_2 \leq C||u - u_h||_0. \tag{2.22}$$

We have

$$
\begin{aligned}
||u-u_h||_0^2 &= \int_\Omega (u-u_h)\,\mathrm{div}(\mathrm{cof}\, D^2 u) Dw\, dx \\
&= \int_\Omega [(\mathrm{cof}\, D^2 u) Dw] \cdot D(u-u_h)\, dx \\
&= \int_\Omega [(\mathrm{cof}\, D^2 u)(Dw - DQw)] \cdot D(u-u_h)\, dx \\
&\qquad + \int_\Omega [(\mathrm{cof}\, D^2 u) DQw] \cdot D(u-u_h)\, dx.
\end{aligned} \tag{2.23}
$$

Recall that for $v_h \in V_0^h$

$$
\begin{aligned}
\int_\Omega \det D^2 u v_h\, dx &= \int_\Omega [(\mathrm{cof}\, D^2 u) Du] \cdot Dv_h\, dx = \int_\Omega f v_h\, dx, \\
\int_\Omega \det D^2 u_h v_h\, dx &= \int_\Omega [(\mathrm{cof}\, D^2 u_h) Du_h] \cdot Dv_h\, dx = \int_\Omega f v_h\, dx.
\end{aligned}
$$

Hence by Lemma 2.4

$$
\begin{aligned}
0 = \int_\Omega \det D^2 u v_h\, dx - \int_\Omega \det D^2 u_h v_h\, dx &= \int_0^1 \bigg\{ \int_\Omega [(\mathrm{cof}\, (1-t) D^2 u_h + t D^2 u) \\
&\qquad (Du - Du_h)] \cdot Dv_h\, dx \bigg\}\, dt, \\
&= \int_0^1 \bigg\{ \int_\Omega [(\mathrm{cof}\, (1-t)(D^2 u_h - D^2 u) + D^2 u) \\
&\qquad (Du - Du_h)] \cdot Dv_h\, dx \bigg\}\, dt.
\end{aligned}
$$

Subtracting from (2.23), using $v_h = Qw$, we obtain

$$
||u-u_h||_0^2 \leq C||w-Qw||_1 ||u-u_h||_1 + +C||u-u_h||_{2,n}^{n-1} ||u-u_h||_2 ||Qw||_2.
$$

Since $||w-Qw||_1 \leq Ch||w||_2$, $||w||_1 \leq ||w||_2$, $||Qw||_2 \leq ||w||_2$, by (2.22),

$$
||u-u_h||_0 \leq Ch^{l+1}|u|_{l+1} + C||u-u_h||_{2,n}^{n-1} h^{l-1} |u|_{l+1}, \tag{2.24}
$$

which gives when $n=2$ since $l \geq 3$, $||u-u_h||_0 = O(h^{l+1})$.
For $n=3$, as in the proof of the H^1 error estimate using ρ_1, we have

$$
\begin{aligned}
||u-u_h||_{2,3} \leq ||u-Qu||_{2,3} + ||Qu-u_h||_{2,3} &\leq h^{l-1}|u|_{l+1,3} + \frac{C}{h^{\frac{4}{3}}} ||Qu - Tu_h||_1 \\
&\leq h^{l-1}|u|_{l+1,3} + Ch^{l-\frac{4}{3}} ||u||_{2,\infty}^2 |u|_{l+1}.
\end{aligned}
$$

Again, here for h sufficiently small, we get from (2.24) that the term h^{l+1} dominates for $n=3$. □

Assume that u_h, the unique strictly convex solution of the discrete Monge-Ampère equation, satisfies $||u-u_h||_1 < \delta_h/2$, where $\delta_h > 0$ has the property that if $||v_h - u||_1 <$

$\delta_h, v_h \in V^h$, then v_h is strictly convex. This is possible by Lemmas 2.5 and 2.6 under the assumption $u \in W^{l+1,\infty}(\Omega), 1+n/2 < l \leq d$. We put

$$X^h = \{\, v_h \in V^h, v_h = g_h \text{ on } \partial\Omega, ||v_h - u_h||_1 < \delta_h/2 \,\},$$

so that for $v_h \in X^h$, v_h is strictly convex. Moreover (2.15) holds.

We consider the mapping $F_h : X^h \to (V_0^h)'$, where $(V_0^h)'$ is the dual space of V_0^h and V_0^h is equipped with $||.||_2$, defined by $\langle F_h(v_h), \psi_h \rangle = \int_\Omega \det D^2 v_h \, \psi_h \, dx, v_h \in V^h, \psi_h \in V_0^h$ and recall that $f > 0$ is continuous. Since Ω is bounded, by L^2 duality, $X^h \subset (V_0^h)'$. We use the notation $||.||$ for the operator norm of an element of a dual space. With these notation, (2.8) can be written $\det D^2 u = f$ in V_0' and (2.10) can be written $F_h(u_h) = f$ in $(V_0^h)'$. We have

Lemma 2.7. *Discrete coercivity: $||F_h'(v_h)(p)|| \geq ch||p||_1$, for all $p \in V_0^h$ and $v_h \in X^h$ for a constant $c > 0$.*

Generalized Lipschitz continuity: $||F_h'(v_h) - F_h'(w_h)|| \leq c(h)||v_h - w_h||_{2,n}, v_h, w_h \in X^h$ where $c(h)$ is independent of h for $n = 2$ and $c(h) = Ch^{-3/2}(\delta_h + ||u_h||_1)$ for $n = 3$. The constant $c(h)$ depends on δ_h and u_h.

Proof. Note that

$$m||p||_1^2 \leq \langle F'(v_h)(p), p \rangle \leq M||p||_1^2, p \in H_0^1(\Omega), \tag{2.25}$$

$v_h \in X^h$ and $||F_h'(v)(p)|| = \sup_{\psi \neq 0} |\langle F_h'(v)(p), \psi \rangle| / ||\psi||_2 \geq |\langle F_h'(v)(p), p \rangle| / ||p||_2 \geq m(||p||_1/||p||_2)||p||_1$. By (2.7), $||p||_1/||p||_2 \geq ch$ for all $p \in V_0^h$. This proves that $||F_h'(v)(p)|| \geq ch||p||_1, p \in V_0^h$ for a constant $c > 0$.

For $v_h, w_h \in X^h$, $\psi \in X^h$, $\eta \in V_0^h$, we have

$$\begin{aligned}
\langle F_h'(v_h)(\psi), \eta \rangle - \langle F_h'(w_h)(\psi), \eta \rangle &= \int_\Omega (\operatorname{div}(\operatorname{cof} D^2 v_h) D\psi) \eta \, dx - \int_\Omega (\operatorname{div}(\operatorname{cof} D^2 w_h) D\psi) \eta \, dx \\
&= -\int_\Omega [(\operatorname{cof} D^2 v_h) D\psi] \cdot D\eta \, dx + \int_\Omega [(\operatorname{cof} D^2 w_h) D\psi] \cdot D\eta \, dx \\
&= \int_\Omega [(\operatorname{cof} D^2 w_h - \operatorname{cof} D^2 v_h) D\psi] \cdot D\eta \, dx
\end{aligned}$$

For $n = 2$ $\operatorname{cof} D^2 w_h - \operatorname{cof} D^2 v_h = \operatorname{cof} D^2(w_h - v_h)$ and as for (2.9), we have

$$\begin{aligned}
|\langle F_h'(v_h)(\psi), \eta \rangle - \langle F_h'(w_h)(\psi), \eta \rangle| &\leq C||v_h - w_h||_2 ||\psi||_2 ||\eta||_2 \\
||F_h'(v_h)(\psi) - F_h'(w_h)(\psi)|| &\leq C||v_h - w_h||_2 ||\psi||_2 \\
||F_h'(v_h) - F_h'(w_h)|| &\leq C||v_h - w_h||_2.
\end{aligned}$$

For $n = 3$, the integral $\int_\Omega [(\operatorname{cof} D^2 w_h - \operatorname{cof} D^2 v_h) D\psi] \cdot D\eta \, dx$ is the sum of terms similar to

$$\int_\Omega \left[\left(\frac{\partial^2 w_h}{\partial y^2} \frac{\partial^2 w_h}{\partial z^2} - \left(\frac{\partial^2 w_h}{\partial y \partial z} \right)^2 \right) - \left(\frac{\partial^2 w_h}{\partial y^2} \frac{\partial^2 w_h}{\partial z^2} - \left(\frac{\partial^2 w_h}{\partial y \partial z} \right)^2 \right) \right] \frac{\partial \psi}{\partial x} \frac{\partial \eta}{\partial x} dx.$$

By the mean value theorem and as with (2.9), and using inverse estimates, we obtain

$$\begin{aligned}|\langle F_h'(v_h)(\psi),\eta\rangle - \langle F_h'(w_h)(\psi),\eta\rangle| &\leq C(||w_h||_{2,3}+||v_h||_{2,3})||v_h-w_h||_{2,3}||\psi||_2||\eta||_2\\ ||F_h'(v_h)(\psi)-F_h'(w_h)(\psi)|| &\leq C(||w_h||_{2,3}+||v_h||_{2,3})||v_h-w_h||_{2,3}||\psi||_2\\ ||F_h'(v_h)-F_h'(w_h)|| &\leq C(||w_h||_{2,3}+||v_h||_{2,3})||v_h-w_h||_{2,3}\\ &\leq Ch^{-\frac{3}{2}}(\delta_h+||u_h||_1)||v_h-w_h||_{2,3},\end{aligned}$$

which proves the result. □

2.2.2. *Convergence of Newton's method for C^1 conforming approximations.* When (1.1) has a smooth solution, the appropriate way to solve the nonlinear equations is to use Newton's method. We establish here that the discretization of the iterates converge to the unique local solution of (2.10). We have

Theorem 2.8. *Under the assumptions of Theorem 2.6, the sequence defined by*

$$F_h'(u_k)(u_{k+1}-u_k) = -(F_h(u_k)-f),$$

with a suitable initial guess converges to the unique convex solution u_h of (2.10) *with a quadratic convergence rate.*

Proof. Since $||F_h'(u_k)(p)|| \geq ch||p||_1, p \in V_0^h$ and $u_{k+1}-u_k \in V_0^h$, the Newton's iterates are well defined. We first show that for $u_k \in X^h$,

$$||F_h(u_{k+1})-f|| \leq c(h)||F_h(u_k)-f||^2. \tag{2.26}$$

Then we establish that if $u_k \in X^h$ $u_{k+1} \in X^h$ $\delta(h)$ sufficiently small. Finally we prove the convergence rate.

Put $p_k = u_{k+1}-u_k$ and let $\psi_h \in V_0^h$. We have by the definition of Newton's method, $F_h(u_h)$, by Fubini's theorem and (2.9),

$$\begin{aligned}\langle F_h(u_{k+1})-f,\psi_h\rangle &= \langle F_h(u_{k+1})-F_h(u_k)+F_h(u_k)-f,\psi_h\rangle\\ &= \langle \int_0^1 F_h'(u_k+tp_k)p_k\,dt + F_h(u_k)-f,\psi_h\rangle\\ &= \langle F_h(u_k)-f+F_h'(u_k)p_k+\int_0^1 F_h'(u_k+tp_k)p_k - F_h'(u_k)p_k\,dt,\psi_h\rangle\\ &\leq \int_0^1 |\langle F_h'(u_k+tp_k)p_k - F_h'(u_k)p_k\,dt,\psi\rangle \leq c(h)||p_k||_{2,n}||p_k||_2||\psi_h||_2,\end{aligned}$$

which implies $||F_h(u_{k+1})-f|| \leq c(h)||p_k||_{2,n}||p_k||_2$. Since $u_k \in X^h$ is strictly convex and $F_h'(u_k)(p_k) = -(F_h(u_k)-f)$, we have $||F_h(u_k)-f|| = ||F_h'(u_k)(p_k)|| \geq ch||p_k||_1$. And by (2.7), $||v_h||_{2,n} \leq ch^{1-n/2}||v_h||_2$ and $||v_h||_2 \leq c/h||v_h||_1$. So $||v_h||_{2,n} \leq ch^{-n/2}||v_h||_1$. That is, $||p_k||_{2,n}||p_k||_2 \leq Ch^{-n/2-1}||p_k||_1^2$. Hence (2.26) holds. Next, we prove that if $u_k \in X^h$, then $u_{k+1} \in X^h$. We have

$$F_h'(u_k)(u_{k+1}-u_h) + F_h'(u_k)(u_h-u_k) = -(F_h(u_k)-f).$$

Hence, using $F_h(u_h) = f$ and the Mean Value Theorem,

$$\begin{aligned}F_h'(u_k)(u_{k+1}-u_h) &= -F_h'(u_k)(u_h-u_k) - F_h'(u_h+\theta(u_k-u_h))(u_k-u_h)\\ &= [F_h'(u_k)-F_h'(u_h+\theta(u_k-u_h))](u_k-u_h), \theta\in[0,1].\end{aligned}$$

We conclude by Lemma 2.7

$$ch||u_{k+1}-u_h||_1 \leq c(h)||u_k-u_h-\theta(u_k-u_h)||_{2,n}||u_k-u_h||_2 \leq c(h)h^{-n/2-1}||u_k-u_h||_1^2.$$

Hence

$$||u_{k+1}-u_h||_1 \leq c(h)||u_k-u_h||_1^2 \leq c(h)\delta_h||u_k-u_h||_1.$$

This gives the convergence rate and shows that, with $\delta(h)$ sufficiently small, $u_{k+1} \in X^h$ when $u_k \in X^h$.

We now show that u_k converges to u_h. Let us assume that u_0 is chosen so that $u_0 \in X^h$ and $u_k \in X^h$ for all k. From (2.26), we obtain

$$\begin{aligned}||F_h(u_k)-f|| &\leq c(h)\cdot c(h)^2 \cdot c(h)^{2^2} \cdots c(h)^{2^{k-1}}||F_h(u_0)-f||^{2^k} \\ &= c(h)^{\frac{1-2^k}{1-2}}||F_h(u_0)-f||^{2^k} = c(h)^{-1}\bigg(c(h)||F_h(u_0)-f||\bigg)^{2^k} = c(h)^{-1}q^{2^k},\end{aligned}$$

where $q=c(h)||F_h(u_0)-f||$.

We now further make the assumption that u_0 is chosen so that $q<1$. Then, by the Mean Value Theorem, $||F_h(u_k)-f|| = ||F_h(u_k)-F_h(u_h)|| = F_h'(v_h)(u_k-u_h)$ for some v_h in X^h. We then have by Lemma 2.7

$$||u_k-u_h||_1 \leq \frac{1}{ch}||F_h(u_k)-f|| \leq \frac{1}{ch}c(h)^{-1}\,q^{2^k},$$

which proves the convergence.

□

2.2.3. *Convergence of the pseudo transient methods for C^1 conforming approximations.* We are now in position to prove the convergence of the iterative methods (1.2). The discretization of (1.2) depends on the choice of L: Given $\nu>0$ and a suitable initial guess, find $u_{k+1} \in V^h, u_{k+1}=g$ on $\partial\Omega$ such that we have for all $\psi_h \in V_0^h$, when L is the Laplace operator

(2.27)

$$\begin{aligned}-\nu\int_\Omega (Du_{k+1}-Du_k)\cdot D\psi_h\,dx + \langle F_h'(u_k)(u_{k+1}-u_k),\psi_h\rangle \\ = \langle -(F_h(u_k)-f),\psi_h\rangle,\end{aligned}$$

and when L is the negative of the identity,

$$-\nu\int_\Omega (u_{k+1}-u_k)\psi_h\,dx + \langle F_h'(u_k)(u_{k+1}-u_k),\psi_h\rangle = \langle -(F_h(u_k)-f),\psi_h\rangle. \tag{2.28}$$

Above, we have made the abuse of notation of denoting by both u_k the solution of the iterative methods at both the continuous and discrete level. In the remainder of this paper, only discrete solutions are considered. This alleviates the notation.

Theorem 2.9. *Let Ω be convex with a Lipschitz continuous boundary and assume that the spaces $V^h=S_d^r(\mathcal{T})$ have the optimal approximation property* (2.5) *and satisfy the inverse estimates* (2.7)*. A sequence defined by either* (2.27) *or* (2.28) *with a suitable initial guess and a suitable value of ν converges to the unique strictly convex solution of* (2.10) *for h sufficiently small. Moreover the convergence rate is linear.*

Proof. Define $\mathcal{M}_i : V_0^h \to (V_0^h)', i = 1, 2$ for $v, \psi_h \in V_0^h$ by

$$\langle \mathcal{M}_1(v), \psi_h \rangle = \int_\Omega Dv \cdot D\psi_h \, dx, \; \langle \mathcal{M}_2(v), \psi_h \rangle = \int_\Omega v\psi_h \, dx.$$

We note that

$$||\mathcal{M}_i(v)|| \leq ||v||_1, v \in V_0^h, i = 1, 2. \tag{2.29}$$

Next, for $p \in V_0^h$, we have by the inverse inequality

$$\begin{aligned} || - \nu\mathcal{M}_1(p) + F_h'(v_h)(p)|| &= \sup\nolimits_{\psi_h \neq 0} \frac{| - \nu\mathcal{M}_1(p)(\psi_h) + F_h'(v_h)(p)(\psi_h)|}{||\psi_h||_2} \\ &\geq \frac{| - \nu\mathcal{M}_1(p)(p) + F_h'(v_h)(p)(p)|}{||p||_2} \\ &= \frac{| - \nu|p|_1^2 + F_h'(v_h)(p)(p)|}{||p||_2} \geq \nu|p|_1 \frac{|p|_1}{||p||_2} + m||p||_1 \frac{||p||_1}{||p||_2} \\ &\geq (\nu C_1 h + C_3 h)||p||_1. \end{aligned}$$

Similarly

$$|| - \nu\mathcal{M}_2(p) + F_h'(v_h)(p)|| \geq \nu||p||_0 \frac{||p||_0}{||p||_2} + m||p||_1 \frac{||p||_1}{||p||_2} \geq (\nu C_2 h^2 + C_3 h)||p||_1.$$

We therefore have

$$||p||_1 \leq \frac{1}{\nu C_i h^i + C_3 h} || - \nu\mathcal{M}_i(p) + F_h'(v_h)(p)||, p \in V_0^h, i = 1, 2. \tag{2.30}$$

We can now determine under which conditions when $u_k \in X^h$ we have $u_{k+1} \in X^h$ as well. Using $F_h(u_h) = f$ and the Mean Value Theorem,

$$\begin{aligned} -\nu\mathcal{M}_i(u_{k+1} - u_h) + F_h'(u_k)(u_{k+1} - u_h) &= -\nu\mathcal{M}_i(u_k - u_h) + F_h'(u_k)(u_k - u_h) \\ &\quad - F_h'(u_h + \theta(u_k - u_h))(u_k - u_h) \\ &= [F_h'(u_k) - F_h'(u_h + \theta(u_k - u_h))](u_k - u_h) \\ &\quad - \nu\mathcal{M}_i(u_k - u_h), \theta \in [0, 1]. \end{aligned}$$

Using (2.30) and (2.29), the generalized Lipschitz continuity property of F_h' and inverse inequalities, we get

$$\begin{aligned} (\nu C_i h^i + C_3 h)||u_{k+1} - u_h||_1 &\leq \nu||u_{k+1} - u_h||_1 + c(h)||u_{k+1} - u_h||_{2,n}||u_{k+1} - u_h||_2 \\ &\leq \nu||u_k - u_h||_1 + c(h)||u_k - u_k||_1^2 \\ &\leq (\nu + c(h)\delta(h))||u_k - u_h||_1. \end{aligned}$$

We therefore have

$$||u_{k+1} - u_h||_1 \leq \frac{\nu + c(h)\delta(h)}{\nu C_i h^i + C_3 h} ||u_k - u_h||_1. \tag{2.31}$$

We note that $(\nu + c(h)\delta(h))/(\nu C_i h^i + C_3 h) < 1$ is equivalent to $\nu(1 - C_i h^i) < C_3 h - c(h)\delta(h)$. This shows that if we choose h sufficiently small so that $1 - C_i h^i > 0$, and then $\delta(h)$ sufficiently small so that $C_3 h - c(h)\delta(h) > 0$ and then ν sufficiently small and possibly $\delta(h)$ smaller, $u_{k+1} \in X^h$ when $u_k \in X^h$.

We now assume that u_0 is chosen in X^h. We have for $i=1,2$

$$\begin{aligned}\langle F_h(u_{k+1})-f,\psi_h\rangle &= \langle F_h(u_{k+1})-F_h(u_k)+F_h(u_k)-f,\psi_h\rangle\\ &= \langle F_h(u_{k+1})-F_h(u_k),\psi_h\rangle - \langle \nu\mathcal{M}_i(u_{k+1}-u_k),\psi_h\rangle\\ &\qquad - \langle F_h'(u_k)(u_{k+1}-u_k),\psi_h\rangle\\ &= \langle \int_0^1 [F_h'(u_k+t(u_{k+1}-u_k))-F_h'(u_k)](u_{k+1}-u_k)\,dt\\ &\qquad - \langle \nu\mathcal{M}_i(u_{k+1}-u_k),\psi_h\rangle.\end{aligned}$$

We conclude that

$$||F_h(u_{k+1})-f|| \leq c(h)||u_{k+1}-u_k||_{2,n}||u_{k+1}-u_k||_2 + \nu||\mathcal{M}_i(u_{k+1}-u_k)||.$$

By (2.7), $||u_{k+1}-u_k||_{2,n}||u_{k+1}-u_k||_2 \leq c(h)||u_{k+1}-u_k||_1^2$. Moreover by (2.29), $||\mathcal{M}_i(u_{k+1}-u_k)|| \leq ||u_{k+1}-u_k||_1$, $i=1,2$. We therefore have

$$||F_h(u_{k+1})-f|| \leq c(h)||u_{k+1}-u_k||_1^2 + \nu||u_{k+1}-u_k||_1.$$

Finally, by (2.30) and the definition of the iterative methods (2.27) and (2.28), $||u_k - u_h||_1 \leq \frac{1}{\nu C_i h^i + C_3 h}||F_h(u_k)-f||$, $i=1,2$. We conclude that

$$||F_h(u_{k+1})-f|| \leq \frac{c(h)}{(\nu C_i h^i + C_3 h)^2}||F_h(u_k)-f||^2 + \frac{\nu}{\nu C_i h^i + C_3 h}||F_h(u_k)-f||$$

Put $\alpha(\nu,h,n)=c(h)/(\nu C_i h^i + C_3 h)^2$ and $\beta(\nu,h,n)=\nu/(\nu C_i h^i + C_3 h)$. We note that $\beta(\nu,h,n)=\nu/(\nu C_i h^i + C_3 h)<1$ is equivalent to $\nu(2-C_i h^i)<C_3 h$. We assume that h is chosen sufficiently small so that $(2-C_i h^i)>0$ and ν is chosen sufficiently small so that $\beta(\nu,h,n)<1/2$. Let $q=||F_h(u_0)-f||$ and assume that u_0 is chosen so that $\alpha(\nu,h,n)q<1/2$ and $q<1$. We then have

$$s \equiv \alpha(\nu,h,n)q + \beta(\nu,h,n) < 1.$$

It follows that

$$||F_h(u_1)-f|| \leq \alpha(\nu,h,n)||F_h(u_0)-f||^2 + \beta(\nu,h,n)||F_h(u_0)-f|| = sq,$$

and since $s<1$,

$$\begin{aligned}||F_h(u_2)-f|| &\leq ||F_h(u_1)-f||(\alpha(\nu,h,n)||F_h(u_1)-f|| + \beta(\nu,h,n))\\ &\leq ||F_h(u_1)-f||s \leq s^2 q.\end{aligned}$$

We conclude that $||F_h(u_k)-f|| \leq s^k q$. Using $F_h(u_h)=f$, the Mean Value Theorem and the discrete coercivity property of F_h', we have

$$||u_k-u_h||_1 \leq \frac{1}{ch}||F_h(u_k)-f|| \leq \frac{q}{ch}s^k,$$

from which the convergence follows. The convergence rate is given by (2.31). □

Remark 2.10. *The analysis above does not indicate whether* (2.27) *should be preferred over* (2.28). *We view* (2.27) *as a preconditioned version of* (2.28). *Moreover, the numerical results indicate that the use of the Laplacian preconditioner improves the convexity property of the numerical solution.*

2.2.4. *Convergence of the time marching methods for C^1 conforming approximations.* We now turn to the proof of one of the main results of this paper, the convergence analysis of the iterative method (1.3) for the Monge-Ampère equation.

By Poincare's inequality, $|w|_1 \leq ||w||_1 \leq C_4|w|_1, w \in H_0^1(\Omega)$. Let $\gamma_h = \delta_h/C_4$. We now assume that u_h the unique strictly convex solution of the discrete Monge-Ampère equation satisfies $|u-u_h|_1 < \gamma_h/2$ and we now define X^h as

$$X^h = \{\, v_h \in V^h, v_h = g_h \,\text{on}\, \partial\Omega, |v_h - u_h|_1 < \gamma_h/2 \,\}.$$

Then for $v_h \in X^h$, by Lemma 2.5, $m|w|_1^2 \leq \int_\Omega [(\operatorname{cof} D^2 v_h) Dw] \cdot Dw \; dx \leq C_4^2 M |w|_1^2$, $w \in H_0^1(\Omega)$. Without loss of generality, relabeling $C_4^2 M$ as M, we will assume that for $v_h \in X^h$,

$$m|w|_1^2 \leq \int_\Omega [(\operatorname{cof} D^2 v_h) Dw] \cdot Dw \; dx \leq M|w|_1^2, w \in H_0^1(\Omega). \tag{2.32}$$

Let $\nu = (M+m)/2$ and define a mapping $T_1 : X^h \to (V_0^h)'$ by

$$\langle T_1(v_h), \psi_h \rangle = \int_\Omega Dv_h \cdot D\psi_h \, dx + \frac{1}{\nu} \int_\Omega (\det D^2 v_h - f)\, \psi_h \, dx, v_h \in X^h, \psi_h \in V_0^h. \tag{2.33}$$

The following lemma will make it possible to show that T_1 is a strict contraction.

Lemma 2.11. *For $v_h \in X^h$,*

$$||T_1'(v_h)||_* \equiv sup_{\psi_h \in V_0^h, \psi_h \neq 0} \frac{||T_1'(v_h)(\psi_h)||}{|\psi_h|_1} \leq sup_{\psi_h \in V_0^h, \psi_h \neq 0} \frac{|T_1'(v_h)(\psi_h)(\psi_h)|}{|\psi_h|_1^2} \leq \frac{M-m}{M+m}.$$

Proof. Let $\alpha = \sup_{\psi_h \in V_0^h, \psi_h \neq 0} \frac{|T_1'(v_h)(\psi_h)(\psi_h)|}{|\psi_h|_1^2}$. We have

$$|T_1'(v_h)(\psi_h)(\psi_h)| \leq \alpha |\psi_h|_1^2, \psi_h \in V_0^h. \tag{2.34}$$

Since for $\mu_h \in V_0^h$, $||T_1'(v_h)(\mu_h)|| = \sup_{\eta_h \in V_0^h, \eta_h \neq 0} |T_1'(v_h)(\mu_h)(\eta_h)|/|\eta_h|_1$, we obtain

$$||T_1'(v_h)||_* = \sup_{\mu_h, \eta_h \in V_0^h, \mu_h, \eta_h \neq 0} \frac{|T_1'(v_h)(\mu_h)(\eta_h)|}{|\mu_h|_1 |\eta_h|_1}.$$

But

$$\begin{aligned} T_1'(v_h)(\mu_h)(\eta_h) &= \int_\Omega D\mu_h \cdot D\eta_h \, dx - \frac{1}{\nu} \int_\Omega [(\operatorname{cof} D^2 v_h) D\mu_h] \cdot D\eta_h \; dx \\ &= \int_\Omega [(I - \frac{1}{\nu} (\operatorname{cof} D^2 v_h)) D\mu_h] \cdot D\eta_h \; dx, \end{aligned}$$

where I denote the $n \times n$ identity matrix. Hence

$$\frac{|T_1'(v_h)(\mu_h)(\eta_h)|}{|\mu_h|_1 |\eta_h|_1} = \int_\Omega [(I - \frac{1}{\nu} (\operatorname{cof} D^2 v_h)) D \frac{\mu_h}{|\mu_h|_1}] \cdot D \frac{\eta_h}{|\eta_h|_1} \; dx.$$

Next, we note that for fixed $v_h \in X^h$, we can define a bilinear form on V_0^h by the formula

$$(p, q) = \int_\Omega [(I - \frac{1}{\nu} (\operatorname{cof} D^2 v_h)) Dp] \cdot Dq \; dx.$$

Then since

$$(p, q) = \frac{1}{4} ((p+q, p+q) - (p-q, p-q)),$$

we obtain

$$||T_1'(v_h)||_* = \sup_{\mu_h,\eta_h \in V_0^h, \mu_h,\eta_h \neq 0} \frac{1}{4} \bigg| \int_\Omega [(I - \frac{1}{\nu}\mathrm{cof}\, D^2 v_h) D(\frac{\mu_h}{|\mu_h|_1} + \frac{\eta_h}{|\eta_h|_1}] \cdot D(\frac{\mu_h}{|\mu_h|_1} + \frac{\eta_h}{|\eta_h|_1})\ dx$$

$$- \int_\Omega [(I - \frac{1}{\nu}\mathrm{cof}\, D^2 v_h) D(\frac{\mu_h}{|\mu_h|_1} - \frac{\eta_h}{|\eta_h|_1}] \cdot D(\frac{\mu_h}{|\mu_h|_1} - \frac{\eta_h}{|\eta_h|_1})\ dx \bigg|$$

$$\leq \frac{\alpha}{4}\left(\left| \frac{\mu_h}{|\mu_h|_1} + \frac{\eta_h}{|\eta_h|_1} \right|_1^2 + \left| \frac{\mu_h}{|\mu_h|_1} - \frac{\eta_h}{|\eta_h|_1} \right|_1^2 \right) = \alpha.$$

Using (2.15), we have

$$(1 - \frac{M}{\nu})|w|_1^2 \leq \int_\Omega [(I - \frac{1}{\nu}(\mathrm{cof}\, D^2 v_h)) Dw] \cdot Dw\ dx \leq (1 - \frac{m}{\nu})|w|_1^2, w \in H_0^1(\Omega).$$

Since $\nu = (M+m)/2$, $1-M/\nu = -(M-m)/(M+m)$ and $1-m/\nu = (M-m)/(M+m)$, we conclude that $\alpha \leq (M-m)/(M+m)$. □

We can now prove the following lemma

Lemma 2.12. *The mapping T_1 is a strict contraction in X^h with contraction constant $(M-m)/(M+m)$ for $\nu = (M+m)/2$.*

Proof. Let v_h and $w_h \in X^h$. Then, using the mean value theorem

$$||T_1(w_h) - T_1(v_h)|| = || \int_0^1 T_1'(v_h + t(w_h - v_h))(w_h - v_h)\, dt||$$
$$\leq \int_0^1 ||T_1'(v_h + t(w_h - v_h))(w_h - v_h)||\, dt.$$

Since $w_h - v_h \in V_0^h$ and $v_h + t(w_h - v_h) \in X^h, t \in [0,1]$, we obtain by Lemma 2.11,

$$||T_1(w_h) - T_1(v_h)|| \leq \int_0^1 \frac{M-m}{M+m}|w_h - v_h|_1\, dt = \frac{M-m}{M+m}|w_h - v_h|_1.$$

□

Remark 2.13. *For the operator T_1 to be a strict contraction, it is enough to have ν sufficiently large. From the proofs of Lemmas 2.11 and 2.12, we need to have max $(|1-m/\nu|, |1-M/\nu|) < 1$. This is equivalent to $\nu >$ max $(m, M/2)$ or $M/2 < \nu \leq m$.*

Remark 2.14. *By the inverse inequality, we have*

$$C_5 h^2 |w_h|_1^2 \leq ||w_h||_0^2 \leq C_6 |w_h|_1^2, w_h \in V_0^h.$$

It follows that

$$(C_5 h^2 - \frac{M}{\nu})|w_h|_1^2 \leq \int_\Omega w_h^2\, dx + \int_\Omega [\frac{1}{\nu}(\mathrm{cof}\, D^2 v_h) D w_h] \cdot D w_h\ dx \leq (C_6 - \frac{m}{\nu})|w_h|_1^2, w_h \in V_0^h.$$

As in the proofs of Lemmas 2.11 and 2.12, we conclude that for max $(|C_5 h^2 - \frac{M}{\nu}|, |C_6 - \frac{m}{\nu}|) < 1$, the mapping $T_2 : X^h \to (V_0^h)'$ defined by

$$(2.35) \quad < T_2(v_h), \psi_h \rangle = \int_\Omega v_h \psi_h\, dx + \frac{1}{\nu} \int_\Omega (\det D^2 v_h - f)\, \psi_h\, dx, v_h \in X^h, \psi_h \in V_0^h.$$

is a strict contraction. But $|C_5 h^2 - \frac{M}{\nu}| < 1$ *is equivalent to* $(\nu < M/(C_5 h^2))$ *or* $(\nu \geq M/(C_5 h^2)$ *and* $h < 1/\sqrt{C_5})$. *And similarly,* $|C_6 - \frac{m}{\nu}| < 1$ *is equivalent to* $(\nu < m/(C_6))$ *or* $(\nu \geq m/(C_6)$ *and* $1 < 1/C_6)$. *We conclude that* T_2 *is a strict contraction for* ν *sufficiently small or for* h *sufficiently small and* ν *in an appropriate range.*

We can now claim our main result, which is the convergence to u_h of the sequence defined by $u_{k+1} \in V^h, u_{k+1} = g_h$ on $\partial\Omega$ and

$$\nu \int_\Omega Du_{k+1} \cdot D\psi_h \, dx = \nu \int_\Omega Du_k \cdot D\psi_h \, dx + \int_\Omega (\det D^2 u_k - f) \, \psi_h \, dx, \psi_h \in V_0^h. \tag{2.36}$$

Theorem 2.15. *Let* Ω *be convex with a Lipschitz continuous boundary and assume that the spaces* $V^h = S_d^r(\mathcal{T})$ *have the optimal approximation property* (2.5) *and satisfy the inverse estimates* (2.7)*. The sequence defined by* (2.36) *converges to the unique strictly convex solution* u_h *of* (2.10) *for any initial guess* u_0 *in* X^h *and a suitable* $\nu > 0$ *with a linear convergence rate.*

Proof. The proof parallels Theorem 5.4 in [19]. Let us assume first that $u_k \in X^h$. We have using (2.10), or equivalently $\det D^2 u_h = f$ in $(V_0^h)'$,

$$\begin{aligned} \int_\Omega D(u_{k+1} - u_h) \cdot D\psi_h \, dx &= \int_\Omega D(u_k - u_h) \cdot D\psi_h \, dx + \frac{1}{\nu} \int_\Omega \det D^2 u_k \, \psi_h \, dx \\ &\qquad - \frac{1}{\nu} \int_\Omega \det D^2 u_h \, \psi_h \, dx \\ &= \langle T_1(u_k) - T_1(u_h), \psi_h \rangle. \end{aligned}$$

Taking $\psi_h = u_{k+1} - u_h$, we obtain

$$|u_{k+1} - u_h|_1^2 \leq ||T_1(u_k) - T_1(u_h)|| \, |u_{k+1} - u_h|_1 \leq \frac{M-m}{M+m} |u_k - u_h|_1 |u_{k+1} - u_h|_1,$$

where for simplicity, we assume that the finite dimensional V_0^h is equipped with the $| \, |_1$ norm of $H_0^1(\Omega)$. We conclude that

$$|u_{k+1} - u_h|_1 \leq \frac{M-m}{M+m} |u_k - u_h|_1.$$

This also shows that if $u_k \in X^h$, then $u_{k+1} \in X^h$ and concludes the proof. $\square$

Remark 2.16. *It follows from the above result and Remark 2.14 that for a suitable initial guess and a suitable* $\nu > 0$*, the sequence defined by*

$$\nu \int_\Omega u_{k+1} \psi_h \, dx = \nu \int_\Omega u_k \psi_h \, dx + \int_\Omega (\det D^2 u_k - f) \, \psi_h \, dx, \psi_h \in V_0^h, \tag{2.37}$$

with $u_{k+1} \in V^h, u_{k+1} = g_h$ *on* $\partial\Omega$ *also converges to the unique strictly convex solution of 2.10. Obviously the convergence properties of* (2.36) *and* (2.37) *depend on the contraction constants of* T_1 *and* T_2 *respectively. Thus* (2.36) *is more robust than* (2.37) *since the choice of* ν *for* (2.36) *is less dependent on the discretization parameter* h*. As suggested in* [26] *in the context of monotone schemes, the use of the Laplacian preconditioner results in a more efficient algorithm.*

3. Numerical results

We give a brief review of the spline element method and give numerical results for both the latter and standard finite difference methods. The choice of the latter emphasizes the broad range of applications of the methods we proposed. The numerical results with finite difference methods complete the results with the spline element method in the degenerate case $f > 0$ in Ω and in three dimension. In that case, essentially at the discrete level $f \geq c_0 > 0$ for a constant c_0 as only non zero values are used. In general we did not try to choose the value of ν that would give the smallest number of iterations except in Tables 4 and 5 where we compare the performance of the two methods.

3.1. **Spline element discretization.** We refer to [1, 5, 6, 7, 28, 3] for a description of the spline element method. We describe the method for linear problems and recall that the problems (1.3) are linear problems. Let $u \in V = H_0^m(\Omega), m \geq 1$ solve a variational problem $a(u,v) = f(v)$ with the conditions of the Lax-Milgram lemma satisfied. Take V_h as the spline space $S_d^r(\mathcal{T})$ of smoothness r and degree d, (2.4). For $r=0$ and $d=1$ we have the space of piecewise linear continuous functions.

First, start with a representation of a piecewise discontinuous polynomial as a vector in $\mathbb{R}^N$, for some integer $N > 0$. Then express boundary conditions and constraints including global continuity or smoothness conditions as linear relations. In our work, we use the Bernstein basis representation, [1, 3] which is very convenient to express smoothness conditions and very popular in computer aided geometric design. Hence the term "spline" in the name of the method. We can therefore identify the space V_h with $\{c \in \mathbb{R}^N, Rc = G\}$ for some integer N, matrix R and vector G. The discrete problem consists in finding $c \in V_h, c^T K d = F^T d$ for all $d \in V_h$ for a suitable stiffness matrix K and a load vector F. Introducing a Lagrange multiplier λ, the functional

$$K(c)d - L^T d + \lambda^T R d,$$

vanishes identically on V_h. The stronger condition

$$K(c) + \lambda^T R = L^T,$$

along with the side condition $Rc = G$ are the discrete equations to be solved. We are lead to saddle point problems

$$\begin{pmatrix} K & R^T \\ R & 0 \end{pmatrix} \begin{pmatrix} \mathbf{c} \\ \lambda \end{pmatrix} = \begin{bmatrix} F \\ G \end{bmatrix}.$$

The ellipticity assures uniqueness of the component c and the saddle point problems are solved by a version of the augmented Lagrangian algorithm

$$(K + \frac{1}{\mu} R^T R) c^{(l+1)} = K^T c^{(l)} + \frac{1}{\mu} R^T G, \quad l = 1, 2, \ldots \tag{3.1}$$

The convergence properties of the iterative method were given in [2]. Extensive implementation details can be found in [1, 6].

h	n_{it}	L^2 norm	rate	H^1 norm	rate	H^2 norm	rate
$1/2^1$	236	$4.1569\ 10^{-6}$		$6.5142\ 10^{-5}$		$1.9364\ 10^{-3}$	
$1/2^2$	233	$1.1504\ 10^{-7}$	5.17	$2.3915\ 10^{-6}$	4.77	$1.3444\ 10^{-4}$	3.85
$1/2^3$	233	$3.2406\ 10^{-9}$	5.15	$8.4120\ 10^{-8}$	4.83	$8.9366\ 10^{-6}$	3.92
$1/2^4$	233	$4.5857\ 10^{-10}$	2.82	$4.7246\ 10^{-9}$	4.15	$6.0706\ 10^{-7}$	3.88

TABLE 1. Time marching method for Test 1, S^1_5, $\nu = 50$

d	n_{it}	L^2 norm	H^1 norm	H^2 norm
3	1	$1.2338\ 10^{-2}$	$7.6984\ 10^{-2}$	$4.4411\ 10^{-1}$
4	270	$1.6289\ 10^{-3}$	$1.4719\ 10^{-2}$	$1.3983\ 10^{-1}$
5	135	$1.5333\ 10^{-3}$	$8.7312\ 10^{-3}$	$6.0412\ 10^{-2}$
6	424	$1.2491\ 10^{-4}$	$9.7458\ 10^{-4}$	$1.0473\ 10^{-2}$
Rate		$0.18\ 0.25^{d-1}$	$4.57\ 0.25^{d}$	$60.85\ 0.3^{d+1}$

TABLE 2. Time marching method for Test 2 (3D) on $\mathcal{I}_1$, $\nu = 50$

d	n_{it}	L^2 norm	H^1 norm	H^2 norm
3	1	$3.1739\ 10^{-3}$	$2.3005\ 10^{-2}$	$2.4496\ 10^{-1}$
4	651	$3.2385\ 10^{-4}$	$3.5599\ 10^{-3}$	$5.2262\ 10^{-2}$
5	744	$2.2730\ 10^{-5}$	$3.8977\ 10^{-4}$	$8.8978\ 10^{-3}$
6	652	$1.1956\ 10^{-6}$	$2.2056\ 10^{-5}$	$6.0437\ 10^{-4}$
Rate		$0.72\ 0.072^{d-1}$	$29.44\ 0.1^{d}$	$861.43\ 0.14^{d+1}$

TABLE 3. Time marching method for Test 2 (3D) on $\mathcal{T}_2$, $\nu = 50$

3.2. **Numerical results with the spline element method.** Unless otherwise indicated, all numerical simulations below are for $r = 1$ and the domain is $[0,1]^n, n = 2,3$. For $n = 2$, the computational domain is the unit square $[0,1]^2$ which is first divided into squares of side length h. Then each square is divided into two triangles by the diagonal with negative slope. For $n = 3$, the initial tetrahedral partition $\mathcal{T}_1$ consists in six tetrahedra. Each tetrahedron is then uniformly refined into 8 subtetrahedra forming $\mathcal{T}_2$. In the tables, n_{it} denotes the number of iterations. We refer to [1, 6] for implementation details of the method.

We use test functions suggested in [16, 9, 20].

Test 1: $u(x,y) = e^{(x^2+y^2)/2}$ so that $f(x,y) = (1 + x^2 + y^2)e^{(x^2+y^2)}$ and $g(x,y) = e^{(x^2+y^2)/2}$ on $\partial\Omega$.

Test 2: $u(x,y,z) = e^{(x^2+y^2+z^2)/3}$ so that $f(x,y,z) = 8/81(3+2(x^2+y^2+z^2)e^{(x^2+y^2+z^2)}$ and $g(x,y,z) = e^{(x^2+y^2+z^2)/3}$ on $\partial\Omega$.

Barring roundoff errors, the methods introduced in this paper capture smooth solutions. For the two dimensional test function, Test 1, we give numerical results for successive refinements and for the three dimensional test function, we give numerical results for increasing values of the degree d on two successive refinements.

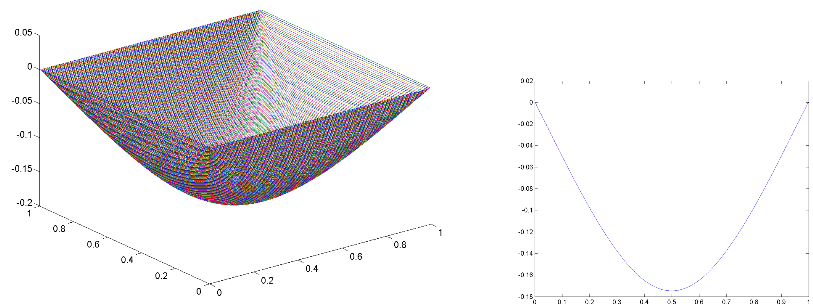

FIGURE 1. Pseudo transient with $L = \Delta$, Test 3, convex solution: $h = 1/2^4, d = 5, \nu = 7.5$.

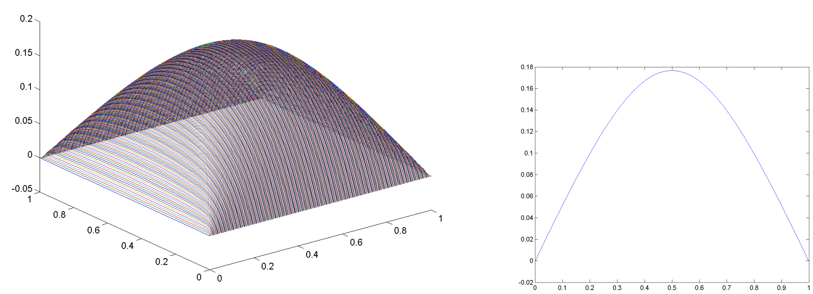

FIGURE 2. Time marching, Test 3, concave solution: $h = 1/2^4, d = 5, \nu = 50$.

In the context of approximations by finite dimensional spaces, many methods proposed, [16, 9, 20], fail to fully capture the convexity of the solution on the test case

Test 3: $g(x, y) = 0$ and $f(x, y) = 1$.

The pseudo transient method with $L = \Delta$ for plane problems enforces element by element $\Delta u \geq 0$ and for a non degenerate problem $\det D^2 u = f > 0$. This implies that the numerical solution has Hessian positive definite element by element, which when combined with C^1 continuity gives numerical convexity [30], Lemma 1. The numerical results are given in Figure 4 with a plot of the graph of the solution along the line $y = x$.

For the same test case, there is a concave solution. The concavity property of the concave solution obtained with the time marching method are better than the one obtained by the vanishing moment methodology, [20]. This is illustrated in Figure 2.

We now discuss how the two methods compare. First, we are solving the same discrete equations (2.10) by different iterative methods. Second, we noticed that the smaller ν, the smaller the number of iterations. Thus for a smooth solution, the correct value of ν to take in the pseudo transient method is $\nu = 0$ which is exactly Newton's

h	ν	n_{it}	time	L^2 norm	rate
$1/2^1$	0	6	3.032810^{+0}	2.195410^{-2}	
$1/2^2$	0	5	8.136510^{+0}	3.609710^{-3}	2.60
$1/2^3$	0	6	3.823010^{+1}	1.068510^{-3}	1.76
$1/2^4$	3	56	1.597910^{+3}	3.766610^{-4}	1.50

TABLE 4. Pseudo-transient method with $L = \Delta$, Test 4 $r = 1, d = 3$

h	ν	n_{it}	time	L^2 norm	rate
$1/2^1$	2	35	$6.2191\ 10^{+0}$	2.072110^{-2}	
$1/2^2$	2	89	$6.0553\ 10^{+1}$	1.857910^{-3}	3.48
$1/2^3$	4.5	64	$1.6849\ 10^{+2}$	5.043810^{-4}	1.88
$1/2^4$	11.5	151	1.703810^{+3}	2.113210^{-4}	1.25

TABLE 5. Time marching method with $L = \Delta$, Test 4 $r = 1, d = 3$

method. In fact, Newton's method has been shown to have a quadratic convergence rate while the pseudo transient methods and time marching methods are shown in Theorems 2.9 and 2.15 to have a linear convergence rate. Moreover the numerical errors of Tables 1, 2 and 3 are essentially the ones obtained with Newton's method as expected. We compare the performance of the methods on a non-smooth solution with known solution.

Test 4: $u(x, y) = -\sqrt{2 - x^2 - y^2}$ with corresponding f and g.

The time listed is in seconds and obtained on an imac running Mac OS 10.6.8 with a 2.4 Ghz intel core 2 duo and 4 GB of SDRAM memory. While for small values of h the time marching method appears to take significantly more time, it is also significantly more accurate. For $h = 1/2^4$ the time took by the two methods is almost the same with the time marching method giving a more accurate solution.

Note that the convergence analysis in Hölder spaces require the domain to be smooth while the convergence analysis in Sobolev spaces only assumes the existence of a smooth solution with the domain allowed to have Lipschitz continuous boundary. We conclude this section with a test problem on a non square domain to contrast with results that can be obtained with the finite difference methods.

Test 5: we consider the unit circle discretized with a Delanauy triangulation with 824 triangles and $u(x, y) = x^2 + y^2 - 1$ which vanishes on the boundary, Figure 3.

3.3. **Numerical results with finite difference methods.** We used a standard finite difference discretization of (1.3). The results are given on Tables 6, 7 and Figure 4. Numerical errors are in the maximum norm. The method also behaves well, in the non degenerate case, for the three dimensional analogues of these cases (also used in [23]).

For $u(x, y, z) = -\sqrt{3 - x^2 - y^2 - z^2}$, it was necessary to stabilize the method using (1.4) with m large. With $m = 25, \nu = 50$, we obtained the same results with the

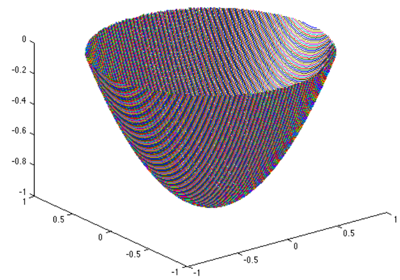

FIGURE 3. $u(x,y) = x^2 + y^2 - 1$ on a non square domain with pseudo transient $\nu = 0, r = 1, d = 3$

	h				
ν	$1/2^2$	$1/2^3$	$1/2^4$	$1/2^5$	$1/2^6$
50	$3.9093\ 10^{-3}$	$1.0340\ 10^{-3}$	$2.6643\ 10^{-4}$	$6.6964\ 10^{-5}$	$1.6781\ 10^{-5}$

TABLE 6. Smooth solution $u(x,y) = e^{(x^2+y^2)/2}$

	h				
ν	$1/2^4$	$1/2^5$	$1/2^6$	$1/2^7$	$1/2^8$
15	2.211310^{-2}	$1.5945\ 10^{-1}$	$3.8944\ 10^{0}$	$6.8634\ 10^{+1}$	$1.1695\ 10^{+3}$
150	$2.2113\ 10^{-2}$	$1.6920\ 10^{-2}$	$1.2440\ 10^{-2}$	$3.3583\ 10^{-2}$	$1.1316\ 10^{0}$
250	$2.2113\ 10^{-2}$	$1.6920\ 10^{-2}$	$1.2440\ 10^{-2}$	$8.9702\ 10^{-3}$	$2.1435\ 10^{-1}$
600					$1.6745\ 10^{-2}$
800					$6.4054\ 10^{-3}$

TABLE 7. Non smooth solution (not in $H^2(\Omega)$) $u(x,y) = -\sqrt{2 - x^2 - y^2}$

three dimensional iterative method introduced in [4],

$$\Delta u_{k+1} = ((\Delta u_k)^3 + 9(f - \det D^2 u_k))^{\frac{1}{3}}.$$

The errors in the maximum norm were given by $3.0976\,10^{-3}$, $1.0432\,10^{-3}$, $1.4169\,10^{-3}$, $1.3766\,10^{-3}$, $1.1017\,10^{-3}$, $8.3671\,10^{-4}$, $3.6635\,10^{-5}$ for $h = 1/2^k, k = 2, \ldots, 8$ respectively.

The convergence rate for the smooth solution is 2, and 0.50 for the two dimensional non smooth solution, and 0.25 for the three dimensional non smooth solution. For the last one, we took only in account errors for $h = 1/2^k, k = 4, 5, 8$.

Acknowledgments

The author acknowledges discussions with F. Celiker, B. Cockburn, W. Gangbo, R. Glowinski, M.J. Lai, R. Nochetto, A. Oberman and A. Regev. The author was supported in part by NSF grant DMS-0811052 and the Sloan Foundation. This research was supported in part by the Institute for Mathematics and its Applications with funds provided by the National Science Foundation.

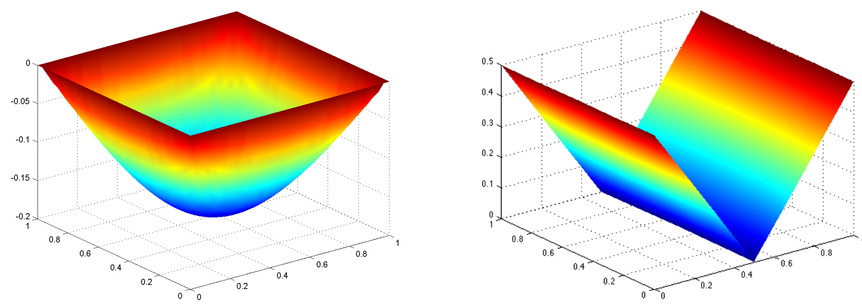

FIGURE 4. No known exact solution (left), $f(x, y) = 1, g(x, y) = 0, h = 1/2^5, \nu = 15$ and $u(x, y) = |x - 1/2|$ with $g(x, y) = |x - 1/2|$ and $f(x, y) = 0$, $h = 1/2^2, \nu = 5$

References

1. G. Awanou, *Energy methods in 3D spline approximations of the Navier-Stokes equations*, Ph.D. Dissertation, University of Georgia, Athens, Ga, 2003.
2. G. M. Awanou and M. J. Lai, *On convergence rate of the augmented Lagrangian algorithm for nonsymmetric saddle point problems*, Appl. Numer. Math. **54** (2005), no. 2, 122–134.
3. Gerard Awanou, *Robustness of a spline element method with constraints*, J. Sci. Comput. **36** (2008), no. 3, 421–432.
4. ———, *Spline element method for the Monge-Ampère equation*, Submitted, 2012.
5. Gerard Awanou and Ming-Jun Lai, *Trivariate spline approximations of 3D Navier-Stokes equations*, Math. Comp. **74** (2005), no. 250, 585–601 (electronic).
6. Gerard Awanou, Ming-Jun Lai, and Paul Wenston, *The multivariate spline method for scattered data fitting and numerical solution of partial differential equations*, Wavelets and splines: Athens 2005, Mod. Methods Math., Nashboro Press, Brentwood, TN, 2006, pp. 24–74.
7. Victoria Baramidze and Ming-Jun Lai, *Spherical spline solution to a PDE on the sphere*, Wavelets and splines: Athens 2005, Mod. Methods Math., Nashboro Press, Brentwood, TN, 2006, pp. 75–92.
8. Jean-David Benamou and Yann Brenier, *A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem*, Numer. Math. **84** (2000), no. 3, 375–393.
9. Jean-David Benamou, Brittany D. Froese, and Adam M. Oberman, *Two numerical methods for the elliptic Monge-Ampère equation*, M2AN Math. Model. Numer. Anal. **44** (2010), no. 4, 737–758.
10. Klaus Böhmer, *On finite element methods for fully nonlinear elliptic equations of second order*, SIAM J. Numer. Anal. **46** (2008), no. 3, 1212–1249.
11. Susanne C. Brenner, Thirupathi Gudi, Michael Neilan, and Li-yeng Sung, *C^0 penalty methods for the fully nonlinear Monge-Ampère equation*, Math. Comp. **80** (2011), no. 276, 1979–1995.
12. Susanne C. Brenner and L. Ridgway Scott, *The mathematical theory of finite element methods*, second ed., Texts in Applied Mathematics, vol. 15, Springer-Verlag, New York, 2002.
13. R. Courant and D. Hilbert, *Methods of mathematical physics. Vol. II*, Wiley Classics Library, John Wiley & Sons Inc., New York, 1989, Partial differential equations, Reprint of the 1962 original, A Wiley-Interscience Publication.
14. O Davydov and A. Saeed, *Numerical solution of fully nonlinear elliptic equations by Bohmer's method*, Preprint.
15. ———, *Stable splitting of bivariate spline spaces by bernstein-bézier methods*, To appear in LNCS volume of the Proceedings of the 7th Conference on Curves and Surfaces, held in Avignon, France, June 24-30, 2010.
16. E. J. Dean and R. Glowinski, *Numerical methods for fully nonlinear elliptic equations of the Monge-Ampère type*, Comput. Methods Appl. Mech. Engrg. **195** (2006), no. 13-16, 1344–1386.

17. Bradley W. Dyer and Don Hong, *Algorithm for optimal triangulations in scattered data representation and implementation*, J. Comput. Anal. Appl. **5** (2003), no. 1, 25–43, Approximation theory and wavelets (Austin, TX, 1999).
18. Lawrence C. Evans, *Partial differential equations*, Graduate Studies in Mathematics, vol. 19, American Mathematical Society, Providence, RI, 1998.
19. I. Faragó and J. Karátson, *Numerical solution of nonlinear elliptic problems via preconditioning operators: theory and applications*, Advances in Computation: Theory and Practice, vol. 11, Nova Science Publishers Inc., Hauppauge, NY, 2002.
20. X. Feng and M. Neilan, *Vanishing moment method and moment solutions for second order fully nonlinear partial differential equations*, J. Sci. Comput. **38** (2009), no. 1, 74–98.
21. Xiaobing Feng and Michael Neilan, *Analysis of Galerkin methods for the fully nonlinear Monge-Ampère equation*, J. Sci. Comput. **47** (2011), no. 3, 303–327.
22. B. D. Froese and A. M. Oberman, *Convergent filtered schemes for the monge ampere partial differential equation*, Submitted, 2012.
23. ———, *Fast finite difference solvers for singular solutions of the elliptic Monge-Ampère equation*, J. Comput. Phys. **230** (2011), no. 3, 818–834.
24. B.D. Froese and A.M. Oberman, *Convergent finite difference solvers for viscosity solutions of the elliptic Monge-Ampere equation in dimensions two and higher*, SIAM J. Numer. Anal. **49** (2011), no. 4, 1692–1714.
25. David Gilbarg and Neil S. Trudinger, *Elliptic partial differential equations of second order*, Classics in Mathematics, Springer-Verlag, Berlin, 2001, Reprint of the 1998 edition.
26. Roland Glowinski, *Numerical methods for fully nonlinear elliptic equations*, ICIAM 07—6th International Congress on Industrial and Applied Mathematics, Eur. Math. Soc., Zürich, 2009, pp. 155–192.
27. Gary Harris and Clyde Martin, *The roots of a polynomial vary continuously as a function of the coefficients*, Proc. Amer. Math. Soc. **100** (1987), no. 2, 390–392.
28. Xian-Liang Hu, Dan-Fu Han, and Ming-Jun Lai, *Bivariate splines of various degrees for numerical solution of partial differential equations*, SIAM J. Sci. Comput. **29** (2007), no. 3, 1338–1354 (electronic).
29. C. T. Kelley and David E. Keyes, *Convergence analysis of pseudo-transient continuation*, SIAM J. Numer. Anal. **35** (1998), no. 2, 508–523. MR 1618838 (99d:65261)
30. Ming-Jun Lai, *Convex preserving scattered data interpolation using bivariate C^1 cubic splines*, J. Comput. Appl. Math. **119** (2000), no. 1-2, 249–258, Dedicated to Professor Larry L. Schumaker on the occasion of his 60th birthday.
31. Ming-Jun Lai and Larry L. Schumaker, *On the approximation power of bivariate splines*, Adv. Comput. Math. **9** (1998), no. 3-4, 251–279.
32. ———, *Spline functions on triangulations*, Encyclopedia of Mathematics and its Applications, vol. 110, Cambridge University Press, Cambridge, 2007.
33. Omar Lakkis and Tristan Pryer, *A nonvariational finite element method for nonlinear elliptic problems*, Submitted, 2011.
34. Grégoire Loeper and Francesca Rapetti, *Numerical solution of the Monge-Ampère equation by a Newton's algorithm*, C. R. Math. Acad. Sci. Paris **340** (2005), no. 4, 319–324.
35. Michael Neilan, *A unified analysis of some finite element methods for the Monge-Ampère equation*, Submitted.
36. Adam M. Oberman, *Wide stencil finite difference schemes for the elliptic Monge-Ampère equation and functions of the eigenvalues of the Hessian*, Discrete Contin. Dyn. Syst. Ser. B **10** (2008), no. 1, 221–238.
37. V. I. Oliker and L. D. Prussner, *On the numerical solution of the equation $(\partial^2 z/\partial x^2)(\partial^2 z/\partial y^2) - ((\partial^2 z/\partial x \partial y))^2 = f$ and its discretizations. I*, Numer. Math. **54** (1988), no. 3, 271–293.
38. A. M. Ostrowski, *Solution of equations and systems of equations*, Pure and Applied Mathematics, Vol. IX. Academic Press, New York-London, 1960. MR 0127525 (23 #B571)
39. Roy Plastock, *Homeomorphisms between Banach spaces*, Trans. Amer. Math. Soc. **200** (1974), 169–183.

40. Neil S. Trudinger and Xu-Jia Wang, *Boundary regularity for the Monge-Ampère and affine maximal surface equations*, Ann. of Math. (2) **167** (2008), no. 3, 993–1028.

Department of Mathematics, Statistics, and Computer Science, M/C 249. University of Illinois at Chicago, Chicago, IL 60607-7045, USA

E-mail address: awanou@uic.edu

URL: http://www.math.uic.edu/~awanou